

Result No.	Query			DB	ID	Description
	Score	Match	Length			
C	1	761	7.6	841	6	CD644465
	2	752	7.5	790	8	BZ604860
	3	742.4	7.4	744	8	BZ609651
	4	727.8	7.3	2331	3	AK030249
	5	648.6	6.5	690	1	AL040111
	6	640	6.4	640	4	BM834501
	7	618	6.2	618	5	BX955300
	8	592.4	5.9	595	5	BP944766
	9	581.4	5.8	583	5	BP336988
	10	557.6	5.6	775	8	BZ603781
C	11	543.8	5.4	571	7	CR753532
	12	536	5.4	585	5	BP366185
	13	528.2	5.3	557	8	B16035
	14	516	5.2	598	5	BP270262
	15	498.4	5.0	583	5	BP229047
	16	485.8	4.9	730	8	BZ606973
	17	473.8	4.7	569	2	AW753609
	18	473.2	4.7	895	9	CR094054
	19	418.8	4.2	768	2	BP237807
	20	402.8	4.0	458	5	BP201868
C	21	394	3.9	411	4	BG007233
	22	382	3.8	479	1	A1380386
	23	379.8	3.8	594	7	CV409050
C	24	378.2	3.8	636	6	CB457791
	25	378.2	3.8	636	6	CB457791


```
Db 121 TCGCCGCCCCCTCTGGATGAAATAGGAGCCGTGGATGAGGGAAGTGGTCTCTGAAGACG 180
Qy 4075 GATCTGAGGATGGCTTCCCGAAGGAATCAATCTGGGTAAAGCTGCCCTGTCTCGTCCCG 4134
Db 181 GATCTGAGGATGGCTTCCCGAAGGAATCAATCTGGGTAAAGCTGCCCTGTCTCGTCCCG 240
Qy 4135 TGTCTTCCCGCTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 4194
Db 241 TGTCTTCCCGCTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 296
Qy 4195 CAACGCTGGCCAGGAATGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTGGTAT 4254
Db 297 CAACGCTGGCCAGGAATGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTGGTAT 356
Qy 4255 CACTCTGTGTAAAGTCAATTAACCTCTCAGGCGCTTAATTTTCTCATTTCTGTAATAACAG 4314
Db 357 CACTCTGTGTAAAGTCAATTAACCTCTCAGGCGCTTAATTTTCTCATTTCTGTAATAACAG 416
Qy 4315 GGTGAGTTAAGAGTCTCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 4374
Db 417 GGTGAGTTAAGAGTCTCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 476
Qy 4375 TTTTGTCTCAAAACACACTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAAT 4434
Db 477 TTTTGTCTCAAAACACACTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAAT 536
Qy 4435 GCTTCTTAACCTGGGGGATTTTGTCTCCCAATCAGTATCTGGCAATGTCTGGAGGCAATTT 4494
Db 537 GCTTCTTAACCTGGGGGATTTTGTCTCCCAATCAGTATCTGGCAATGTCTGGAGGCAATTT 596
Qy 4495 GGTGTCTACTCTGTGTGGGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 4553
Db 597 GGTGTCTACTCTGTGTGGGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 656
Qy 4554 CACTGTCTCAGTGTGTACAGTGCACAGGACAGCCCATCATCAAGAAATATCTGTCTCC 4613
Db 657 CACTGTCTCAGTGTGTACAGTGCACAGGACAGCCCATCATCAAGAAATATCTGTCTCC 716
Qy 4614 AATGTCAATAGTTTGTGAGCAATGAGAGACCTAGCTTCACTTAAAGTTTCTTGGCGTTC 4673
Db 717 AATGTCAATAGTTTGTGAGCAATGAGAGACCTAGCTTCACTTAAAGTTTCTTGGCGTTC 775
Qy 4674 CTGATCTTTTCTG 4687
Db 776 CTGATCTTTTCTG 789
```

```
RESULT 3
BZ609651
LOCUS WHADE767P Human MCF7 breast cancer cell line library (MCF7_1) Homo sapiens genomic clone MCF7_1-20M7, genomic survey sequence.
DEFINITION BZ609651.1 GI:31518212
ACCESSION BZ609651
VERSION 1
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 744)
Volik, S., Zhao, S., Chin, K., Brebner, J. H., Herndon, D. R., Tao, Q., Komel, D., Huang, G., Lapuk, A., Kuo, W.-L., Magrane, G., de Jong, P., Gray, J. W. and Collins, C.
End-sequence profiling: Sequence-based analysis of aberrant genomes
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
22709111
12788976
Contact: Volik SV
Colin Collins' lab
UCSF Comprehensive Cancer Center
UCSF Box 0808, San Francisco, CA 94143-0808, USA
Tel: 415 502 7066
```

Fax: 415 502 5665
Email: svolik@cc.ucsf.edu
This clone is available from Amplicon Express
http://www.genomex.com
Class: BAC ends.
Location/Qualifiers
1. 744
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="MCF7_1-20M7"
/sex="female"
/clone_lib="Human MCF7 breast cancer cell line library (MCF7_1)"
/notes="Vector: pECBAC1; Site 1: HindIII; This library was constructed from MCF7 breast cancer cell line by Amplicon Express (http://www.genomex.com) using their standard procedure."

ORIGIN

Query Match 7.4%; Score 742.4; DB 8; Length 744;
Best Local Similarity 99.9%; Pred. No. 2.5e-107;
Matches 743; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 763 AGCTTCTATGTATAGAAACCTAAGTCACCTTCACATTTCTGATTAGCAGAGTAATTGAATAT 822
Db 1 AGCTTCTATGTATAGAAACCTAAGTCACCTTCACATTTCTGATTAGCAGAGTAATTGAATAT 60

Qy 823 TCTTTTCAATGTGTAGCTCTATCCCGAAGAACCAACAGAAATTTGGAAGTGTAAAGGCCATC 882
Db 61 TCTTTTCAATGTGTAGCTCTATCCCGAAGAACCAACAGAAATTTGGAAGTGTAAAGGCCATC 120

Qy 883 CTATAGTTTAAACCAACTGGTAAATAGATATATAGAAAGATGTGGTATGTGGCAGTGACA 942
Db 121 CTATAGTTTAAACCAACTGGTAAATAGATATATAGAAAGATGTGGTATGTGGCAGTGACA 180

Qy 943 ACTTGAAGTTGTGACTAGAACTCGGGTCTCTGGAGTGTCTATTATATACACACCAAGCT 1002
Db 181 ACTTGAAGTTGTGACTAGAACTCGGGTCTCTGGAGTGTCTATTATATACACACCAAGCT 240

Qy 1003 GGTCCACGAGCCATGTGTGATCTCTCCATGTGATAGCAACAAGAAAGACATTCAGGAC 1062
Db 241 GGTCCACGAGCCATGTGTGATCTCTCCATGTGATAGCAACAAGAAAGACATTCAGGAC 300

Qy 1063 ATTCTTTCTTTACCCCTAATCTTGATCTGAGTCTTATTAGAAAAGCTTAAATGTTAAA 1122
Db 301 ATTCTTTCTTTACCCCTAATCTTGATCTGAGTCTTATTAGAAAAGCTTAAATGTTAAA 360

Qy 1123 GATCTAGTTTATTCAAAACTAAAGATAACAAGAGTATGAGAAATTTCTATTTCCGAGTGT 1182
Db 361 GATCTAGTTTATTCAAAACTAAAGATAACAAGAGTATGAGAAATTTCTATTTCCGAGTGT 420

Qy 1183 AAAGGAGGAGATGTTTCTTGGCTTCTGAGGCTGCGAGCCCTTCTTCTGCTTTTAAGGA 1242
Db 421 AAAGGAGGAGATGTTTCTTGGCTTCTGAGGCTGCGAGCCCTTCTTCTGCTTTTAAGGA 480

Qy 1243 AGTAGAGAGAGGAGGAGAAAGTAAAGTATCTTTTGGTGGAAATTAAGGTTTCTTCTCGGAG 1302
Db 481 AGTAGAGAGAGGAGGAGAAAGTAAAGTATCTTTTGGTGGAAATTAAGGTTTCTTCTCGGAG 540

Qy 1303 TAGTTTGCATGCTTTTGGTGGTGGAAATTAAGTAACTGACTTAAGTTTAAAGTAGTT 1362
Db 541 TAGTTTGCATGCTTTTGGTGGTGGAAATTAAGTAACTGACTTAAGTTTAAAGTAGTT 600

Qy 1363 GGGACTATTTAAACCAATGCTTATCCATGTTTGGCATAAAGGAGAGGTTATGGCTT 1422
Db 601 GGGACTATTTAAACCAATGCTTATCCATGTTTGGCATAAAGGAGAGGTTATGGCTT 660

Qy 1423 TAGAAGTTAAATTTCTTCTCCAGGAGTGAATAATTTAGCTTCTTAAACCAAGCAGCAGAGCTA 1482
Db 661 TAGAAGTTAAATTTCTTCTCCAGGAGTGAATAATTTAGCTTCTTAAACCAAGCAGCAGAGCTA 720

Qy 1483 AATAAAGTAATTTTCTCCACCTGGCC 1506


```
QY 8880 GGCTCTTCCACAATTGCCGGCAATTCTTTGAGTAAAGTTTGATTCACAGTATCACCTG 8939
Db 1036 GCGCTCCAGCGCTGTCTGCAATCTCTTTGAGCAAGTGTCTGATCCCGAGCATTCGCTG 1095
QY 8940 TCCATTGTTGACCTTCAAGACATTTTATCCGAAGATTTTATGATGACACAGAGACTGGA 8999
Db 1096 CCCCTTTTGTACTTTCAGAGCTTTTATCCGAAGTCTTTATGATGACACAGAGACTTGA 1155
QY 9000 GCATAAATACAACTCCTGACGTTCTATAAAACTGTGGAACCAAGTCCTTGTTCAGAGTCG 9059
Db 1156 GCACAGGTACACCTGACCGGACACAGACCGGACGACGAGTCTGTGCTGAGAACAG 1215
QY 9060 ACGTACCGGATGCCCGCAGCGTTGCTGGGAAAGATGTGCTCTCCCTCCCTAGTTTCTG 9119
Db 1216 GCGTACCGGCTGCCCTCGGCTTGTCTGGGAAGATGTCCCTCCCTTGTCTGSCCTGCA 1275
QY 9120 TAAACCAAGCCCAAGTCTGTTTCCCGGCGCAGTCAAAATCCCTGCCATCTGCGAAGG 9179
Db 1276 CAAGCCCAAGCCCAAGTCTGTTTCCCAAGTGTGTTTCCCAAGTGTGCGTCTGCTATGGA 1335
QY 9180 GAAGCAGAGCCCTCTCGGCGCAGCAAGGCCCTCTGACTTCAGGGATAGACTCTAGCAC 9239
Db 1336 TCGCAGGGGCGCTCGGGGCGAAGCAAGCACCCAGACTTCAGGACCAAGACAGCAC 1395
QY 9240 TTTAGCCCCAAGTAACCTGAAGTCCCAAGACCAAGCAAGTGTGGGGGTCCAAAGGGGC 9299
Db 1396 TTTAGCCCCAAGTAACCTGAAGTCCCAAGTGTGGGGGTCCAAAGGGGCACAG--- 1452
QY 9300 CGGCACAGGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTGTTCCCTCGACCGGA 9359
Db 1453 TGGCACCAGGCGAGCAGTGTGTTTCCCAAGTGTGTTTCCCAAGTGTGTTTCCCTGCTATGGA 1512
QY 9360 TAAGACAAAAGACCCGAGACAAATTTGAACCTCTTCCAGTAGTCTTCTCAGGCCAC 9419
Db 1513 TAAGTGAAGAGACCTGAGCAAACTGAAGTCTTCCAGAGTCTTCCAGTCTTCCAGTCTTCC 1572
QY 9420 CCTCGGCGAGTAAACATCAATGTTTCCATCGACTACCCCGCCCAAGAAACACAGCCCGTG 9479
Db 1573 CCTCAGCAGTAATAATAGCAACGGTCTGTTGAGTATCCCGTGNAGTTGACGGCCCATG 1632
QY 9480 GGCACCTCCGGGAGAGACTATTTCTGTAATCGGAGTGCAGCAATCTGACAGCAAT 9539
Db 1633 GGCACAGCAAGGAGAGACTACTACTGCCATCGGAATTTCTGGCAGTGCAGCAGAGTA 1692
QY 9540 TGTGAGCCCTTCCAAAAGACTGAAGTCCAGCGTGTGCGCTTGCCTTGCAGCTTACAGCC 9599
Db 1693 CAGTGAAGCCATCCCAAAAGACTCAAGTCCAGTGGCGGTGCTCTGGACACAGAGCATGC 1752
QY 9600 CGGGGCCAATTACAGAGAGGCTATGACCTTCCCAAGTACCATATGTTGTCAGAGGCAATCAC 9659
Db 1753 AGGACCAATGGCAGAGGGGCTTTGAGCTCCCAAGTACCATATGTTGTCAGGAGCATCAC 1812
QY 9660 ATCACTGTTACCGAGAGCTGTGTATCCGTGCGAGGCGCTCCCTCCCAAAACCAAGTT 9719
Db 1813 CTCTTGTCTACCAACAGAGTGTGCGCCCAACCGCTGTGCTGCCCAAAAGCCCGTTT 1872
QY 9720 CCGTACTCCAGGAGTCAATTTCTCCAAATGCTGCTGCTGTCAGAGCCCTATCGTGG 9779
Db 1873 CCGTGAAGCCCTGGGAGGTGGAGTCAACCCAGTGTGTTGGCTGTGTCAGAGCCCTACAGTGC 1932
QY 9780 CTCCGGGCCAATTACACTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 9839
Db 1933 CTCTGGACCCCTGTATACCTGTGGAACCGGTGGACAGGAGGAGGAGGAGGAGGAGGAGGAGG 1992
QY 9840 AGGTA 9844
Db 1993 AGGAA 1997
```

RESULT 5
AL040111
LOCUS

```
DEFINITION DKF2p434C2113_r1_434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION DKF2p434C2113_5', mRNA sequence.
VERSION AL040111
KEYWORDS AL040111.1 GI:5409078
SOURCE EST.
ORGANISM Homo sapiens (human)
REFERENCE 1 (bases 1 to 690)
AUTHORS Koehrer,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE EST (Koehrer, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
si sequence also available.
This clone (DKF2p434C2113) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
1..690
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKF2p434C2113"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/notes="Vector: pSport1; Site_1: NotI; Site_2: SalI"
ORIGIN
Query Match 6.5%; Score 648.6; DB 1; Length 690;
Best Local Similarity 98.0%; Pred. No. 1.7e-92;
Matches 676; Conservative 0; Mismatches 12; Indels 2; Gaps 2;
QY 1056 TCAGGACATCTTCTTCCCTTACCCCTAATCCTTGATCTGAGTCTTATTAGAAAGCTTAA 1115
Db 1 TCNGGACATCTTCTTCCCTTACCCCTAATCCTTGATCTTCACTCTTATTAGAAAGCTTAA 60
QY 1116 TGTTAAAGATCTAGTTTATTCAAACCTAAAGATTAACAAGGATGATGAGAAATTTCTATTTC 1175
Db 61 TGTTAAAGATCTAGTTTATTCAAACCTAAAGATTAACAAGGATGATGAGAAATTTCTATTTC 120
QY 1176 GGAGTGTAAAGGAGGAGATGTTTCTTGGCTTCTCGAGCTTCGAGCCCTTCCTTGTCT 1235
Db 121 GGAGTGTAAAGGAGGAGATGTTTCTTGGCTTCTCGAGCTTCGAGCCCTTCCTTGTCT 180
QY 1236 TTAAGGAATAGAGAGAGGAGGAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 1295
Db 181 TTAAGGAATAGAGAGAGGAGGAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 240
QY 1296 CTGGGAGTATGTCATGCTTTTGGTGTGTTTGGGTGGAAATTAACCTGACTTAAGTTTAA 1355
Db 241 CTGGGAGTATGTCATGCTTTTGGTGTGTTTGGGTGGAAATTAACCTGACTTAAGTTTAA 300
QY 1356 AGTAGTGGGACTATTTAAACCAATGCTATCCAAATGTTTGGCCATAAAGGAGAGGTA 1415
Db 301 AGTAGTGGGACTATTTAGAACCAATGCTATCCAAATGTTTGGCCATAAAGGAGAGGTA 360
QY 1416 TTGGCTTTAGAGTAAATTTCTTCCAGGAGTAAATTTAGCTTCTTAAACCAAGCAGC 1475
Db 361 TTGGCTTTAGAGTAAATTTCTTCCAGGAGTAAATTTAGCTTCTTAAACCAAGCAGC 420
QY 1476 AGAGCTAAATAAGTAAATTTTCCACCTGGCCAGTGCATGATGTGAAGGTAGATTAAAAA 1535
```

Db 421 AGAGCTAAATAAGTAATTTTCCACCTGGCCAGTGCATGATGTAAGAGTAGATTAAAA 480
QY 1536 AATGAGAGGCCCATTTTCTGATGAAGAAGCTAAGCCATGTTGAAACAGCCCTGTTGAGGA 1595
Db 481 AATCAGAGGACCCATTTTCTGATGAAGAAGCTAAGCCATGTTGAAACAGCCCTGTTGAGGA 540
QY 1596 TTTTATTTTAAATCTATACATTCACAAAGGAGCTTTGTGTATGTTCTTCCCTATTGTTG 1655
Db 541 TTTTATTTTAAATCTATACATTCACAAAGGAGCTTTGTGTATGTTCTTCCCTATTGTTG 600
QY 1656 TTTGAGCTAGGAAGGCCACCCAGCTGCTTTGTTCAAGGCAGAAAGTGTGTAAGCAAG-C 1714
Db 601 TTTGAGCTAGGAAGGCCACCCAGCTGCTTTGTTGAGGCAGAAAGTGTGTAAGCAAGC 660
QY 1715 TGGGATTTGAACAG-TGGATTGAGGTTTCG 1743
Db 661 CGGGATTTGAACAGTTGGATTGAGGTTTCG 690

RESULT 6
BM834501
LOCUS
DEFINITION
K-EST0109502 S11SNUI Homo sapiens cDNA clone S11SNUI-63-D04 5',
mRNA sequence.
ACCESSION
BM834501
VERSION
BM834501.1 GI:19190910
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. (bases 1 to 640)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
UNPUBLISHED (2002)
CONTACT: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 63 row: D column: 04
High quality sequence stop: 640.
Location/Qualifiers
1. .640
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S11SNUI-63-D04"
/sex="M"
/tissue_type="Stomach"
/cell_type="Lymphoblast-like"
/cell_line="SNU-1"
/lab_host="Top10f"
/clone_lib="S11SNUI"
/note="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI;
Site 2: XhoI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including SfII
site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized with Superscript II using SfII
oligo-dT primer. After first strand synthesis, RNA was
degraded by NaOH treatment and cDNA was amplified by PCR
reaction. The PCR products were digested with SfII and
cloned into DraIII- digested pME18-FL3 vector. The
obtained cDNA vectors were used for transfection of
competent cells E. coli Top10f, by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

FEATURES
source

ORIGIN
Query Match 6.4%; Score 640; DB 4; Length 640;
Best Local Similarity 100.0%; Pred. No. 4.1e-91;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2489 TGCCTTGTGCTGATTCATATTTGAATCGAGGCAATTCGGAACCTCTATGCGCTTGTGTTG 2548
Db 1 TGCCTTGTGCTGATTCATATTTGAATCGAGGCAATTCGGAACCTCTATGCGCTTGTGTTG 60
QY 2549 GAAAGAACCCAGTGACACCATCACTGAGCTTCTTAAAGTTCGAAGAAGTTAGAGACTAT 2608
Db 61 GAAAGAACCCAGTGACACCATCACTGAGCTTCTTAAAGTTCGAAGAAGTTAGAGACTAT 120
QY 2609 ACACCTTCTTTTGAACCTTTTATAATAATATTTGCTCTGCTGTTTTTGGAAACCCAGGCTGT 2668
Db 121 ACACCTTCTTTTGAACCTTTTATAATAATATTTGCTCTGCTGTTTTTGGAAACCCAGGCTGT 180
QY 2669 TAGAGGGGTGAGTGACAAAGTCTTCAAGTGGGCTTATTCCAACTCCAGAAATTTGCCCAAC 2728
Db 181 TAGAGGGGTGAGTGACAAAGTCTTCAAGTGGGCTTATTCCAACTCCAGAAATTTGCCCAAC 240
QY 2729 GGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTT 2788
Db 241 GGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTT 300
QY 2789 AATGTACATGATGGGCCAGAAAGTATTGGCAGCTCTCTTGGCAGTCCGATCGAGATGGA 2848
Db 301 AATGTACATGATGGGCCAGAAAGTATTGGCAGCTCTCTTGGCAGTCCGATCGAGATGGA 360
QY 2849 GGATGCTTGTCAATGAAAGGACCGCTGTTGCTCCATTCCGAGCTACACAAGAAAAA 2908
Db 361 GGATGCTTGTCAATGAAAGGACCGCTGTTGCTCCATTCCGAGCTACACAAGAAAAA 420
QY 2909 TGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTCGAGCAGACCTTCAC 2968
Db 421 TGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTCGAGCAGACCTTCAC 480
QY 2969 ACATTCAGAGACCTTATAACATGTTCTTANTGCAACACCGGCTTACCCTCTGTGAAC 3028
Db 481 ACATTCAGAGACCTTATAACATGTTCTTANTGCAACACCGGCTTACCCTCTGTGAAC 540
QY 3029 AGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAG 3088
Db 541 AGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAG 600
QY 3089 AGAACCTCCCAAGGAAAGATTGCAAGAAAAATGAATTT 3128
Db 601 AGAACCTCCCAAGGAAAGATTGCAAGAAAAATGAATTT 640

RESULT 7
BX955300
LOCUS
DEFINITION
DKFZp781N0270 r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
DKFZp781N0270 5', mRNA sequence.
ACCESSION
BX955300
VERSION
BX955300.1 GI:43437635
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 618)
Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
Fobo,G., Han,M. and Wiemann,S.
EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., et al.)
UNPUBLISHED (2003)
CONTACT: MIPS

MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No s1 sequence available.

This clone (DKFZp781N0270) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de.

FEATURES

Location/Qualifiers
1..618
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp781N0270"
/dev_stage="adult"
/lab_host="DH10B"
/notes="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN

Query Match 6.2%; Score 618; DB 5; Length 618;
Best Local Similarity 100.0%; Pred. No. 1.2e-87;
Matches 618; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2802 GGCACAGAGTATGGCAGCTCTCTGGCAGTCCGATGGAGATGGAGGATGCTTGTCA 2861
DB 1 GGCACAGAGTATGGCAGCTCTCTGGCAGTCCGATGGAGATGGAGGATGCTTGTCA 60

QY 2862 ATCAAGAGGACCGCTGTGTTCATCCGAGCTACACAGAAAAATGTCATCCAAATC 2921
DB 61 ATCAAGAGGACCGCTGTGTTCATCCGAGCTACACAGAAAAATGTCATCCAAATC 120

QY 2922 GAGGGGTATATGCCCTTGGATTGATGCTGAGGAGGAGCTTCCACATTCAGAGAC 2981
DB 121 GAGGGGTATATGCCCTTGGATTGATGCTGAGGAGGAGCTTCCACATTCAGAGAC 180

QY 2982 CTTAATAAATGCTTTAATGCAACACCGGCTTACCTCTGTGAACACAGAGTCTTCGG 3041
DB 181 CTTAATAAATGCTTTAATGCAACACCGGCTTACCTCTGTGAACACAGAGTCTTCGG 240

QY 3042 GTTGAAGCAGATCTCAGTCGCTGTATTAAGTCAGTGGAGAGGAGCTTCCAAAG 3101
DB 241 GTTGAAGCAGATCTCAGTCGCTGTATTAAGTCAGTGGAGAGGAGCTTCCAAAG 300

QY 3102 GAAAGAAATGCAAGGAAATGAATTTAGCTGTGAGTATGCTGAGGAGGAGCTTCA 3161
DB 301 GAAAGAAATGCAAGGAAATGAATTTAGCTGTGAGTATGCTGAGGAGGAGCTTCA 360

QY 3162 GCTTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3221
DB 361 GCTTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420

QY 3222 ATGTGCGGAGAGATTCAGGAGGCTTGTGTTCTTAAATCAATCAGATGCGGAGAT 3281
DB 421 ATGTGCGGAGAGATTCAGGAGGCTTGTGTTCTTAAATCAATCAGATGCGGAGAT 480

QY 3282 GGCAAATCGGGGCCAGAGCAAACTGCAAGAGGCTTGGAGAGTATGCTGAGCAAGATC 3341
DB 481 GGCAAATCGGGGCCAGAGCAAACTGCAAGAGGCTTGGAGAGTATGCTGAGCAAGATC 540

QY 3342 AACAGGTCGTTCAGGTGCAACCGGCGGAGAGATCTCTCTCTTACAAAATCTGCATG 3401
DB 541 AACAGGTCGTTCAGGTGCAACCGGCGGAGAGATCTCTCTCTTACAAAATCTGCATG 600

QY 3402 GTTTGGGCTTCTATTT 3419
DB 601 GTTTGGGCTTCTATTT 618

RESULT 8
BF944766 595 bp mRNA linear EST 22-JAN-2001
LOCUS PM3-NN1174-181000-001-c08 NN1174 Homo sapiens cDNA, mRNA sequence.
DEFINITION

ACCESSION

BF944766.1 GI:12362041

VERSION

EST.

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 595)

AUTHORS

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva W. Jr., Zago,M.A., Bordin,S., Costa,F.P.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,

O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

PROC. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=PM3&t2=PM3-NN1174-

181000-001-c08&t3=2000-10-18&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 585.

Location/Qualifiers

1..595

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="NN1174"

/note="Organ: nervous normal; Vector: puc18; Site_1: SmaI;

Site_2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

ORIGIN

Query Match 5.9%; Score 592.4; DB 4; Length 595;
Best Local Similarity 99.8%; Pred. No. 1.4e-83;
Matches 593; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3041 GGTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGTGCAGAACAGCTCCCAA 3100
DB 1 GGTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGTGCAGAACAGCTCCCAA 60

QY 3101 GGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGATGTGGGAGACATTTAGCT 3160
DB 61 GGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGATGTGGGAGACATTTAGCT 120

QY 3161 CGCTTTTGTATGTTGAGATCCATGAGAACACACAAAGATTTCTTCACTTTACGGGTGTA 3220
DB 121 CGCTTTTGTATGTTGAGATCCATGAGAACACACAAAGATTTCTTCACTTTACGGGTGTA 180

QY 3221 CATGTGCGGAAGAAGATTCAAGGAGCTTTGGTTTCTTAAAAATCAGATGCGGACATAA 3280
DB 181 CATGTGCGGAAGAAGATTCAAGGAGCTTTGGTTTCTTAAAAATCAGATGCGGACATAA 240

QY 3281 TCGCAAAATCGGGGGCCAGAGCAAACTGCAGAGGCTTTGGAGAGTAGTCCAGCAACGAT 3340
DB 241 TCGCAAAATCGGGGGCCAGAGCAAACTGCAGAGGCTTTGGAGAGTAGTCCAGCAACGAT 300

```
QY 3341 CAACGAGGTGCTCCAGTGCACGCGCGAGAGCATCTCCTCTCTTACAAAATCTGCAT 3400
Db |||
QY 301 CAACGAGGTGCTCCAGTGCACGCGCGAGAGCATCTCCTCTCTTACAAAATCTGCAT 360
Db |||
QY 3401 GGTGTTGTGGCTTCTTATTTCCAAATAAAGAAAGTCTTAATTGAGCACCAGGAGTGCACAC 3460
Db |||
QY 361 GGTGTTGTGGCTTCTTATTTCCAAATAAAGAAAGTCTTAATTGAGCACCAGGAGTGCACAC 420
QY 3461 CAAAAAAACTGTTTCGGTACAGAGCGCGCAGACAGACTCTCCAAAGAGGAATGCC 3520
Db |||
QY 421 CAAAAAAACTGTTTCGGTACAGAGCGCGCAGACAGACTCTCCAAAGAGGAATGCC 480
QY 3521 GTCTCTGAGGAGGAGTCTTCTCGAGTTGTTCAACTTGAGACCAAAATCTCACCCCTGAAAC 3580
Db |||
QY 481 GTCTCTGAGGAGGAGTCTTCTCGAGTTGTTCAACTTGAGACCAAAATCTCACCCCTGAAAC 540
QY 3581 GGGGAAGAGCCTGTGCAGATGCATCCCTCAGCTCGATCGTTCAGCCCTTCCA 3634
Db |||
QY 541 GGGGAAGAGCCTGTGCAGATGCATCCCTCAGCTCGATCGTTCAGCCCTTCCA 594

RESULT 9
BP336988 583 bp mRNA linear EST 17-SEP-2004
LOCUS BP336988 Sugano cDNA library, coronary artery smooth muscle cell
DEFINITION Homo sapiens cDNA clone SMR07185, mRNA sequence.
ACCESSION BP336988
VERSION BP336988.1 GI:52266581
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
AUTHORS Mizushima-Sugano,J., Nakai,K. and Sugano,S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
JOURNAL Contact: Yutaka Suzuki
COMMENT Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
1..583
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SMR07185"
/tissue_type="coronary artery"
/cell_type="smooth muscle cell"
/clone_lib="Sugano cDNA library, coronary artery smooth
muscle_cell"

ORIGIN
Query Match 5.8%; Score 581.4; DB 5; Length 583;
Best Local Similarity 99.8%; Pred. No. 7.7e-82;
Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8430 TGAAGATGCACATTAATTAACCGCTGACAGTGGCAACCAAAAATTTGAAAAGATTTTGA 8489
Db |||
QY 1 TGAAGATGCACATTAATTAACCGCTGACAGTGGCAACCAAAAATTTGAAAAGATTTTGA 60
Db |||
QY 8490 TGGTGCACAAAGATGTTTACAGGAGTCCACCTGCAAGCAGCTTAAAGGAGATCGCTTCTGT 8549
Db |||
QY 61 TGGTGCCAAAGATGTTTACAGGAGTCCACCTGCAAGCAGCTTAAAGGAGATCGCTTCTGT 120
QY 8550 TTTTCAGATGTTCTGGGAGCGGCTGCTCTCTCACCAGCACACAAGATACAGATTT 8609
Db |||
QY 121 TTTTCAGATGTTCTGGGAGCGGCTGCTCTCTCACCAGCACACAAGATACAGATTT 180
QY 8610 CCATAAAAATGCAGCTGATGACAGTGTCTATAAGTGAATAAAACCCCTACCCCTGCTTA 8669
```

```
Db |||
QY 181 CCATAAAAATGCAGCTGATGACAGTGTGATAAAGTGAATAAAACCCCTACCCCTGCTTA 240
QY 8670 CTGGACCTGTTAAAAAGAGATCAGCAGTTGAAACTCAGGCAATAACCTCATCTGTAG 8729
Db |||
QY 241 CTGGACCTGTTAAAAAGAGATCAGCAGTTGAAACTCAGGCAATAACCTCATCTGTAG 300
QY 8730 AACCAAGCGGATGTTACTCTCTCCGGATGGCAGTACCACCCATACCTTGAAGTTAG 8789
Db |||
QY 301 AACCAAGCGGATGTTACTCTCTCCGGATGGCAGTACCACCCATACCTTGAAGTTAG 360
QY 8790 CCCCACCAAGAGCAAAACGAGACCGCAGCTGACTGCAGATACAGGCCCAAGTGTGATTTG 8849
Db |||
QY 361 CCCCACCAAGAGCAAAACGAGACCGCAGCTGACTGCAGATACAGGCCCAAGTGTGATTTG 420
QY 8850 TCACGAAAAACCTTTAAATTTATCCGTGGGGCTCTTCACAATTCGCCGGCAATTTCTTT 8909
Db |||
QY 421 TCACGAAAAACCTTTAAATTTATCCGTGGGGCTCTTCACAATTCGCCGGCAATTTCTTT 480
QY 8910 GAGTAAAAATTTGATTCCAAGTATCACCTGTCCATTTTGTACCTTCAGACACATTTATCC 8969
Db |||
QY 481 GAGTAAAAATTTGATTCCAAGTATCACCTGTCCATTTTGTACCTTCAGACACATTTATCC 540
QY 8970 AGAAGTTTAAATGATGCACACGAGACTGGAGCATAAATACAAT 9012
Db |||
QY 541 AGAAGTTTAAATGATGCACACGAGACTGGAGCATAAATACAAT 583

RESULT 10
BP603781 775 bp DNA linear GSS 08-JUN-2003
LOCUS WH3AES1TR Human MCF7 breast cancer cell line library (MCF7_1) Homo
DEFINITION sapiens genomic clone MCF7_1-159C23, genomic survey sequence.
ACCESSION BP603781
VERSION BP603781.1 GI:31512243
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,
AUTHORS Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
Gray,J.W. and Collins,C.
End-sequence profiling: Sequence-based analysis of aberrant genomes
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
JOURNAL 22709111
MEDLINE 12788976
PUBMED
COMMENT Contact: Volik SV
Colin Collins' lab
UCSF Comprehensive Cancer Center
UCSF Box 0808, San Francisco, CA 94143-0808, USA
Tel: 415 502 7066
Fax: 415 502 5665
Email: svolik@cc.ucsf.edu
This clone is available from Amplicon Express
http://www.genomex.com
Class: BAC ends.
FEATURES
source
1..775
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="MCF7_1-159C23"
/sex="female"
/clone_lib="Human MCF7 breast cancer cell line library
(MCF7_1)"
/note="Vector: pECBAC1; Site 1: HindIII; This library was
constructed from MCF7 breast cancer cell line by Amplicon
Express (http://www.genomex.com) using their standard
procedure."
```

ORIGIN

Query Match									
Best Local Similarity 5.6%; Score 557.6; DB 8; Length 775;									
Matches 571; Conservative 0; Mismatches 4; Indels 1; Gaps 1;									
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/ .									
Location/Qualifiers									
1. 571									
/organism="Pongo pygmaeus"									
/mol_type="mRNA"									
/db_xref="taxon:9600"									
/clone="DKFZp469L046"									
/tissue_type="kidney"									
/dev_stage="adult"									
/lab_host="DH10B"									
/clone_lib="469 (synonym: pkid1)"									
/notes="Vector: pSport1_Sfi; Site_1: SfiIa; Site_2: SfiIb"									
ORIGIN									
Query Match 5.4%; Score 543.8; DB 7; Length 571;									
Best Local Similarity 97.0%; Pred. No. 6.8e-76;									
Matches 554; Conservative 0; Mismatches 17; Indels 0; Gaps 0;									
2478 AAGAGAACTAATGCTTTGTGCTGATTCAATTTGAATCGAGGCAATTTGGGAACCTGTATG 253									
1 AAGAGAACTAATGCTTTGTGCTGATTCAATTTGAATCGAGGCAATTTGGGAACCTGTATG 60									
2538 CTTTGTGTTGTGAAAGAACGAGTGACACCATCACTGAGCTTCTTAAAGTTTCAAAGAAGT 259									
61 CGGTGTTGTGAAAGAACGAGTGACACCATCACTGAGCTTCTTAAAGTTTCAAAGAAGT 120									
2598 TAGAGACTATACACTTTCTTTTGAACCTTTTATAATAATATTTGCTCTGGTTTTGGAA 265									
121 TAGAGACTATACACTTTCTTTTGAACCTTTTATAATAATATTTGCTCTGGTTTTGGAA 180									
2658 CCAGGGCTTTAGAGGGGTGAGTGACAAAGTCTTACAAGTGGCCTTTATTCCAACTCCAGA 271									
181 CCAGGGCTTTAGAGGGGTGAGTGACAAAGTCTTACAAGTGGCCTTTATTCCAACTCCAGA 240									
2718 AATTGCCAACCGAACTTTGAGATTATATGCAATCGAAAGTGCACAGAAACATGCCAACT 277									
241 AATTGCCAACCGAACTTTGAGATTATATGCAATCGAAAGTGCACAGAAACATGCCAACT 300									
2778 CAATCCCTCTTAATGTACATGATGGCCAGAGTGAATGGCAGCTCTCTTGGCAGTCCG 283									
301 CAGTCCCTCTTAATGTACATGATGGCCAGAGTGAATGGCAGCTCTCTTGGCAGTCCG 360									
2838 ATGGAGATGGAGATGCCCTTGTCAATGAAAGGACCGCTGTTGTTCCATTCCGAGCTACA 289									
361 ATGGAGATGGAGATGCCCTTGTCAATGAAAGGACCGCTGTTGTTCCATTCCGAGCTACG 420									
2898 CAAGAAAAAATGTTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCAATGTTCTGCAGC 295									
421 CAAGAAAAAATGTTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCAATGTTCTGCAGC 480									
2958 CAGACCTTCACACATTCAGAAACACCTTAATAACATGCTTTAATGCAACACCGGCTTACC 301									
481 CAGACCTTCACACATTCAGAAACACCTTAATAACATGCTTTAATGCAACACCGGCTTACC 540									
3018 CTCTGTGAACACGAGCTTCTTCGGGTTGAAG 3048									
541 CTCTGTGAACACGAGCTTCTTCGGGTTGAAG 571									
RESULT 12									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast TIG Homo									
ACCESSION									
BP366185									
VERSION									
BP366185.1 GI:52296435									
KEYWORDS									
EST.									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 585)									
BP366185									
LOCUS									
DEFINITION									
BP366185 Sugano cDNA library, fetal lung fibroblast T									

AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES Location/Qualifiers
source
1..585
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TIR09337"
/tissue_type="lung"
/cell_type="fibroblast"
/cell_line="TIG"
/dev_stage="fetal"
/clone_lib="Sugano cdna library, fetal lung fibroblast
TIG"

ORIGIN
Query Match 5.4%; Score 536; DB 5; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.2e-74;
Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2401 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACATCTTGTCTCAATTTGAAACA 2460
DB 50 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACATCTTGTCTCAATTTGAAACA 109
QY 2461 CTCATTCAATGGAGCAAAAGAGAACTAATGCTTTGTGCTGATCATATTTGAATCGAGC 2520
DB 110 CTCATTCAATGGAGCAAAAGAGAACTAATGCTTTGTGCTGATCATATTTGAATCGAGC 169
QY 2521 ATTGGGAACCTGTATGCTCTTGTGGAAGAACCGAGTGACACCATCAGCTGACCTCC 2580
DB 170 ATTGGGAACCTGTATGCTCTTGTGGAAGAACCGAGTGACACCATCAGCTGACCTCC 229
QY 2581 TAAAGTTTCCGAAGAAGTTAGAGACTATACATCTTCTTTGAACTTTTATAATAATATT 2640
DB 230 TAAAGTTTCCGAAGAAGTTAGAGACTATACATCTTCTTTGAACTTTTATAATAATATT 289
QY 2641 TGTCTGTGTTTTGGAAACCCAGGCTGTTAGAGGGTGAGTGACAAGTCTTACAAGTGGC 2700
DB 290 TGTCTGTGTTTTGGAAACCCAGGCTGTTAGAGGGTGAGTGACAAGTCTTACAAGTGGC 349
QY 2701 CTTATTCCAACTCCGAAATTCGCCAACCGAACTTTGAGATTATATGCAATCGAAAGTGA 2760
DB 350 CTTATTCCAACTCCGAAATTCGCCAACCGAACTTTGAGATTATATGCAATCGAAAGTGA 409
QY 2761 CAGGAAACATGCGCAACTCAATCCCTCTTAATGTACATGATGGCCAGAGTGATTTGGCA 2820
DB 410 CAGGAAACATGCGCAACTCAATCCCTCTTAATGTACATGATGGCCAGAGTGATTTGGCA 469
QY 2821 GCTCTCTTGGCAGTCGATGGAGATGGAGATGCCCTTGTTCATAGAAAGGACCGCTGTTG 2880
DB 470 GCTCTCTTGGCAGTCGATGGAGATGGAGATGCCCTTGTTCATAGAAAGGACCGCTGTTG 529
QY 2881 TTCCATTCGAGCTACACAGAAAAAATGTATCCAAATCGAGGGGTATATGCC 2936
DB 530 TTCCATTCGAGCTACACAGAAAAAATGTATCCAAATCGAGGGGTATATGCC 585

RESULT 13
B16035/c
LOCUS B16035 B16035 557 bp DNA linear GSS 04-JUN-1998
DEFINITION 346L18.TP CIT978SKA1 Homo sapiens genomic clone A-346L18, genomic
survey sequence.
ACCESSION B16035
VERSION B16035.1 GI:2123784
KEYWORDS GSS.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 557)
AUTHORS Adams,M.D., Kelley,J.M., Rounsley,S.R. and Venter,J.C.
TITLE Use of a BAC End Sequence Database for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Classes: BAC ends.

FEATURES Location/Qualifiers
source
1..557
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="A-346L18"
/sex="female"
/cell_type="Fibroblast"
/clone_lib="CIT978SKA1"
/note="Vector: pBAC108L; Site 1: HindIII; Site 2: HindIII;
Caltech Human BAC Library A1"

ORIGIN
Query Match 5.3%; Score 528.2; DB 8; Length 557;
Best Local Similarity 96.8%; Pred. No. 2e-73;
Matches 539; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 520 AAAAAACCAAAAACTCCACTATTTCCTATTGCCACTATTTCGACTATTTCGATTTTAAAAATAAG 579
DB 557 AAAAAACCAAAAACTCCACTATTTCGACTATTTCGACTATTTCGATTTTAAAAATAAG 498
QY 580 CGTATTTTAGCATCTAAAAAGTAGGAAGACCTCAATAAATAGAGTCTTTGTTCTTGGCCA 639
DB 497 CGTATTTTAGCATCTGCAAGTAGGAAGTCTCAATAAATAGAGTCTGTTCTTGGCCA 438
QY 640 GGGAAAAACAGCGTTGTCAGAAATTTGATAAATCTGTTTTCTAGGGTATGCTGTTATTCAG 699
DB 437 GGGAAAAACAGCGTTGTCAGCATGTGTTAACTGTTTCTAGGGTATGCTGTTATTCAG 378
QY 700 TTAAAAACCTTGCCTGGGAGCTAGCATTCAGTAAATACCTTGTGAATAGCAAAATGAAC 759
DB 377 TTAAAAACCTTGCCTGGGAGCTAGCATTCAGTAAATACCTTGTGAATAGCAAAATGAAC 318
QY 760 TTAAGCTTCTATGTATAGAAACCTTAAGTCACTTCATTCATTCGATAGCAGAGTAATTGAA 819
DB 317 TTAAGCTTCTATGTATAGAAACCTTAAGTCACTTCATTCGATAGCAGAGTAATTGAA 258
QY 820 TATTCTTTCAATGTAGTCTTATCCCGAGAACCAAGAAATTTGGAACTGTAAAGGCC 879
DB 257 TAGTCTTTTCACTGTGTAGATCTATCCCGAGAACCAAGAAATTTGGAACTGTAAAGGCC 198
QY 880 ATCCCTATAGTTTAAACCACTGCTTAAATAGATATAGAAAGATGTGGTATGTGGCAGTG 939
DB 197 ATCCCTATAGTTTAAACCACTGCTTAAATAGATATAGAAAGATGTGGTATGTGGCAGTG 138
QY 940 ACAACTTGAAGGTGTGACTAGAACTCGGGTCTCTGGAGTGTCTTATTTATATACACACCA 999
DB 137 ACAACTTGAAGGTGTGACTAGAACTCGGGTCTCTGGAGTGTCTTATTTATATACACACCA 78
QY 1000 GCTGCTCACCAGCCCATGTGTTGATCTCCATTTGATAGCAACAAAGAAAGACTTTCAG 1059
DB 77 GCTGCTCACCAGCCCATGTGTTGATCTCCATTTGATAGCAACAAAGAAAGACTTTCAG 18

QY 1060 GACATCTTCTTCCTTAC 1076
|||||
Db 17 GACACTCTTCTTCCTTAC 1

RESULT 14

BP270262
LOCUS BP270262 Sugano cDNA library, small intestine Homo sapiens EST 17-SEP-2004
DEFINITION clone KAR01734, mRNA sequence.

ACCESSION BP270262
VERSION BP270262.1 GI:52219609
KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 598)

REFERENCE 1 (bases 1 to 598)

AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,

Mizushima-Sugano, J., Nakai, K. and Sugano, S.

TITLE Sequence comparison of human and mouse genes reveals a homologous

block structure in the promoter regions

JOURNAL Genome Res. 14 (9), 1711-1718 (2004)

COMMENT Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: yusuzuki@ims.u-tokyo.ac.jp.

Location/Qualifiers

1. .598

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="KAR01734"

/tissue_type="small intestine"

/clone_lib="Sugano cDNA library, small intestine"

ORIGIN

Query Match 5.2%; Score 516; DB 5; Length 598;

Best Local Similarity 100.0%; Pred. No. 1.7e-71;

Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2401 AGGGTTTGGAAATCCCTTGTCTCCAGGTTCTGGGATGACTTCTTGTCTCAATTTGAACA 2460

|||||

Db 83 AGGGTTTGGAAATCCCTTGTCTCCAGGTTCTGGGATGACTTCTTGTCTCAATTTGAACA 142

|||||

QY 2461 CTCATTCATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGC 2520

|||||

Db 143 CTCATTCATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGC 202

|||||

QY 2521 ATTGGGAACCCCTGTATGCTTTGTGGAAGAACCAAGTACACCATCACTGAGCTTCC 2580

|||||

Db 203 ATTGGGAACCCCTGTATGCTTTGTGGAAGAACCAAGTACACCATCACTGAGCTTCC 262

|||||

QY 2581 TAAAGTTCCGAAGAAGTTAGAGGACTATACATCTTTCTTTTGAACCTTTTATAATAATAT 2640

|||||

Db 263 TAAAGTTCCGAAGAAGTTAGAGGACTATACATCTTTCTTTTGAACCTTTTATAATAATAT 322

|||||

QY 2641 TGCCTCTGGTTTTTGGAAACCCAGGGCTGTTAGAGGGGTGAGTGACAAAGTCTTCAAGTGGC 2700

|||||

Db 323 TGCCTCTGGTTTTTGGAAACCCAGGGCTGTTAGAGGGGTGAGTGACAAAGTCTTCAAGTGGC 382

|||||

QY 2701 CTTATTTCCAACTCCAGAAATTCGCCAACGAACTTTGAGATATATGCAATCGAAGTGA 2760

|||||

Db 383 CTTATTTCCAACTCCAGAAATTCGCCAACGAACTTTGAGATATATGCAATCGAAGTGA 442

|||||

QY 2761 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGATGGCCAGAGTGAATTTGGCA 2820

|||||

Db 443 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGATGGCCAGAGTGAATTTGGCA 502

|||||

QY 2821 GCTCTCTGGAGTCGATGGAGATGCGGATGCTTTGTCAATGAAAGGACCGCTGTTG 2880

|||||

Db 503 GCTCTCTGGAGTCGATGGAGATGCGGATGCTTTGTCAATGAAAGGACCGCTGTTG 562

QY 2881 TTCCATTCCGAGCTACACAGAAAAAATGTCATCC 2916
|||||
Db 563 TTCCATTCCGAGCTACACAGAAAAAATGTCATCC 598

RESULT 15

BP229047
LOCUS BP229047 Sugano cDNA library, fetal brain Homo sapiens EST 15-SEP-2004
DEFINITION FBR01825, mRNA sequence.

ACCESSION BP229047
VERSION BP229047.1 GI:52101957
KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 583)

REFERENCE 1 (bases 1 to 583)

AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,

Mizushima-Sugano, J., Nakai, K. and Sugano, S.

TITLE Sequence comparison of human and mouse genes reveals a homologous

block structure in the promoter regions

JOURNAL Genome Res. 14 (9), 1711-1718 (2004)

COMMENT Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: yusuzuki@ims.u-tokyo.ac.jp.

Location/Qualifiers

1. .583

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="FBR01825"

/tissue_type="brain"

/dev_stage="fetal"

/clone_lib="Sugano cDNA library, fetal brain"

ORIGIN

Query Match 5.0%; Score 498.4; DB 5; Length 583;

Best Local Similarity 99.8%; Pred. No. 1e-68;

Matches 499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2401 AGGGTTTGGAAATCCCTTGTCTCCAGGTTCTGGGATGACTTCTTGTCTCAATTTGAACA 2460

|||||

Db 83 AGGGTTTGGAAATCCCTTGTCTCCAGGTTCTGGGATGACTTCTTGTCTCAATTTGAACA 142

|||||

QY 2461 CTCATTCATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGC 2520

|||||

Db 143 CTCATTCATGGAGACAAAGAGATCTAATGCTTTGTGCTGATTCATATTTGAATCGAGC 202

|||||

QY 2521 ATTGGGAACCCCTGTATGCTTTGTGGAAGAACCAAGTACACCATCACTGAGCTTCC 2580

|||||

Db 203 ATTGGGAACCCCTGTATGCTTTGTGGAAGAACCAAGTACACCATCACTGAGCTTCC 262

|||||

QY 2581 TAAAGTTCCGAAGAAGTTAGAGGACTATACATCTTTCTTTTGAACCTTTTATAATAATAT 2640

|||||

Db 263 TAAAGTTCCGAAGAAGTTAGAGGACTATACATCTTTCTTTTGAACCTTTTATAATAATAT 322

|||||

QY 2641 TGCCTCTGGTTTTTGGAAACCCAGGGCTGTTAGAGGGGTGAGTGACAAAGTCTTCAAGTGGC 2700

|||||

Db 323 TGCCTCTGGTTTTTGGAAACCCAGGGCTGTTAGAGGGGTGAGTGACAAAGTCTTCAAGTGGC 382

|||||

QY 2701 CTTATTTCCAACTCCAGAAATTCGCCAACGAACTTTGAGATATATGCAATCGAAGTGA 2760

|||||

Db 383 CTTATTTCCAACTCCAGAAATTCGCCAACGAACTTTGAGATATATGCAATCGAAGTGA 442

|||||

QY 2761 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGATGGCCAGAGTGAATTTGGCA 2820

|||||

Db 443 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGATGGCCAGAGTGAATTTGGCA 502

|||||

QY 2821 GCTCTCTGGCAGTCCGATGGAGATGCGGATGCTTTGTCAATGAAAGGACCGCTGTTG 2880

```
||||| 503 GCTCTCTTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATGAAGGACCGCTGTG 562
QY 2881 TTCCATTCCGAGCTACACAA 2900
Db 563 TTCCATTCCGAGCTACACAA 582
```

Search completed: August 1, 2005, 01:38:23
Job time : 18949 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 13:54:32 ; Search time 3682 Seconds
(without alignments)
17578.372 Million cell updates/sec

Title: US-08-731-499-9_COPY_1_10000
Perfect score: 10000
Sequence: 1 CCATCATATTCTTATTTT.....ATTCTTAGATACGCAGTGG 10000

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues
Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCRS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10J_PUBCOMB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10000	100.0	10365	8	US-08-731-499-9
2	9980.6	99.8	36022	13	US-10-087-192-1708
3	1639.4	16.4	5632	15	US-10-177-293-505
4	1559.6	15.6	3186	8	US-08-731-499-10
5	1559.6	15.6	5252	13	US-10-087-192-1709
6	1387.2	13.9	1507	8	US-08-731-499-3
7	1006.8	10.1	124990	20	US-10-684-422-156

Sequence 1705, Ap
Sequence 1706, Ap
Sequence 104193,
Sequence 325113,
Sequence 104193,
Sequence 325113,
Sequence 9711, Ap
Sequence 5711, Ap
Sequence 292, App
Sequence 23411, A
Sequence 114, App
Sequence 4, Appli
Sequence 322, App
Sequence 263139,
Sequence 263139,
Sequence 6870, Ap
Sequence 6882, Ap
Sequence 5634, Ap
Sequence 17624, A
Sequence 802, App
Sequence 802, App
Sequence 163893,
Sequence 163893,
Sequence 3, Appli
Sequence 3, Appli
Sequence 585, App
Sequence 1406, Ap
Sequence 1406, Ap
Sequence 262, App
Sequence 28, Appl
Sequence 16, Appl
Sequence 250, App
Sequence 5, Appli
Sequence 93928, A
Sequence 93929, A
Sequence 93928, A

ALIGNMENTS

RESULT 1
US-08-731-499-9
; Sequence 9, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION:
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOWBEL, David
; APPLICANT: KOMMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR
; TITLE OF INVENTION: US98
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,499
; FILING DATE: 16-OCT-1996
; CLASSIFICATION: 435

Db 1741 TCGAATATCAGTGAACCAAAATATATATCATGAGGTTCCCTCGCCCAAGATGAGTGACCAATTC 1800
Qy 1801 TGAGGTGTTAAGTATTTCTTGAATGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 1860
Db 1801 TGAGGTGTTAAGTATTTCTTGAATGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 1860
Qy 1861 ATTTTGTGACCTCTGTATGTGCAAAATCTTAAGGGGGTGTGTGGGCACTTAGATTTCT 1920
Db 1861 ATTTTGTGACCTCTGTATGTGCAAAATCTTAAGGGGGTGTGTGGGCACTTAGATTTCT 1920
Qy 1921 TGGATGCAGATTTGTTTGTATATGAACAATTTTAAATTTGTTTGTATACACTGGATTT 1980
Db 1921 TGGATGCAGATTTGTTTGTATATGAACAATTTTAAATTTGTTTGTATACACTGGATTT 1980
Qy 1981 AAAATAGTTTAAAGTGTTTAAATTTTTCATCTTAAATTTTACAGTTTCTTATAGTCT 2040
Db 1981 AAAATAGTTTAAAGTGTTTAAATTTTTCATCTTAAATTTTACAGTTTCTTATAGTCT 2040
Qy 2041 TTAGATTTAGGAGGCTGTGATGGCATCCATGTGCTATTTAGTGGCAATTTAAATGT 2100
Db 2041 TTAGATTTAGGAGGCTGTGATGGCATCCATGTGCTATTTAGTGGCAATTTAAATGT 2100
Qy 2101 ATTCAGCTGAATTTAACTGATGGAATTTTAACTGGATGAAATTTCTGATTTAAATGTAAGTACTGACT 2160
Db 2101 ATTCAGCTGAATTTAACTGATGGAATTTTAACTGGATGAAATTTCTGATTTAAATGTAAGTACTGACT 2160
Qy 2161 AGCACTGAATTTAACTGATGGAATTTTAACTGGATGAAATTTCTGATTTAAATGTAAGTACTGACT 2220
Db 2161 AGCACTGAATTTAACTGATGGAATTTTAACTGGATGAAATTTCTGATTTAAATGTAAGTACTGACT 2220
Qy 2221 GGATAAAATGCCAATGATTTAACTGAACGACGTTTAACTGATGGAATTTTAACTGATTTAAATGTAAGTACTGACT 2280
Db 2221 GGATAAAATGCCAATGATTTAACTGAACGACGTTTAACTGATGGAATTTTAACTGATTTAAATGTAAGTACTGACT 2280
Qy 2281 TAAAGTGAAGCAATTTGATGGAATTTTAACTGATGGAATTTTAACTGATGGAATTTTAACTGATTTAAATGTAAGTACTGACT 2340
Db 2281 TAAAGTGAAGCAATTTGATGGAATTTTAACTGATGGAATTTTAACTGATGGAATTTTAACTGATTTAAATGTAAGTACTGACT 2340
Qy 2341 ACCCAACACGCTTCTCTCGCTCTGAGTGTAGTAAACGCTTCTGTTTCTGTTTCTCTCT 2400
Db 2341 ACCCAACACGCTTCTCTCGCTCTGAGTGTAGTAAACGCTTCTGTTTCTGTTTCTCTCT 2400
Qy 2401 AGGGTTTGGAAATCCCTTGTCTCCAGTGTCTGGGATGACTTCTTGTCTCAATTTGAAACA 2460
Db 2401 AGGGTTTGGAAATCCCTTGTCTCCAGTGTCTGGGATGACTTCTTGTCTCAATTTGAAACA 2460
Qy 2461 CTCATTTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGC 2520
Db 2461 CTCATTTCAATGGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGC 2520
Qy 2521 ATTTGGAAACCTGTATGCTTGTCTGGAAGAACGAGTGAACCATCACTGAGCTTCC 2580
Db 2521 ATTTGGAAACCTGTATGCTTGTCTGGAAGAACGAGTGAACCATCACTGAGCTTCC 2580
Qy 2581 TAAAGTTCCGAAGATTTAGAGCACTATACATTTCTTTTGAACCTTTTATAATAATTT 2640
Db 2581 TAAAGTTCCGAAGATTTAGAGCACTATACATTTCTTTTGAACCTTTTATAATAATTT 2640
Qy 2641 TGCTCTGGTTTTTGGAAACCGAGGCTGTAGAGGGTGTAGTGAAGTCTTCAAGTGGC 2700
Db 2641 TGCTCTGGTTTTTGGAAACCGAGGCTGTAGAGGGTGTAGTGAAGTCTTCAAGTGGC 2700
Qy 2701 CTTATTTCCAACTCCGAAATTTGCCAACGAACTTTTGAATTTATATGCAATCGAAGTGA 2760
Db 2701 CTTATTTCCAACTCCGAAATTTGCCAACGAACTTTTGAATTTATATGCAATCGAAGTGA 2760
Qy 2761 CAGGAAACATGCCAACTCAATCCCTCTTAATGTATCATGATGGGCGAGAGTGTGGCA 2820
Db 2761 CAGGAAACATGCCAACTCAATCCCTCTTAATGTATCATGATGGGCGAGAGTGTGGCA 2820
Qy 2821 GCTCTCTGGCAGTCCGATGGAGATGCGCTTGTCAATGAAAGGACCGCTGTTG 2880
Db 2821 GCTCTCTGGCAGTCCGATGGAGATGCGCTTGTCAATGAAAGGACCGCTGTTG 2880

Qy 2881 TTCCATTTCCGAGCTACACAAGAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGG 2940
Db 2881 TTCCATTTCCGAGCTACACAAGAAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGG 2940
Qy 2941 ATTGCAATGTTCTCGAGCCAGACCTTTCACATTTCAGAAGACCTTTAAATAAATGTCCTTAA 3000
Db 2941 ATTGCAATGTTCTCGAGCCAGACCTTTCACATTTCAGAAGACCTTTAAATAAATGTCCTTAA 3000
Qy 3001 TGCACAACCGGCTACCTCTGTGAACCCAGCAGTCTTTCGGGTTGAAGCAGAGTATCTCA 3060
Db 3001 TGCACAACCGGCTACCTCTGTGTGAACCCAGCAGTCTTTCGGGTTGAAGCAGAGTATCTCA 3060
Qy 3061 GTCGGCTTGATAAAAGTCAAGTCCGAACAGAACTTCCCAAGGAAAGAAATTTGAAGGAAA 3120
Db 3061 GTCGGCTTGATAAAAGTCAAGTCCGAACAGAACTTCCCAAGGAAAGAAATTTGAAGGAAA 3120
Qy 3121 ATGAATTTAGCTGTGAGGTATGTTGGCAGACATTTTAGAGTCGCTTTTGTGATGATCC 3180
Db 3121 ATGAATTTAGCTGTGAGGTATGTTGGCAGACATTTTAGAGTCGCTTTTGTGATGATCC 3180
Qy 3181 ACATGAGAACACACAAGATTTCTTTCACTTACGGGTGTAAACATGTGCGGAAGAGATTCA 3240
Db 3181 ACATGAGAACACACAAGATTTCTTTCACTTACGGGTGTAAACATGTGCGGAAGAGATTCA 3240
Qy 3241 AGGAGCTTTGGTTTCTTAAATAATCACATGCGGACACATTAATGCAAAATCGGGGCCAGAA 3300
Db 3241 AGGAGCTTTGGTTTCTTAAATAATCACATGCGGACACATTAATGCAAAATCGGGGCCAGAA 3300
Qy 3301 GCAAACTGCAGCAAGGCTTTGGAGAGTAGTCCAGCAACGATCAACGAGGTGCTCCAGGTGC 3360
Db 3301 GCAAACTGCAGCAAGGCTTTGGAGAGTAGTCCAGCAACGATCAACGAGGTGCTCCAGGTGC 3360
Qy 3361 AGCGGCCGAGAGCATCTCTCTCTTACAAAATCTGCAATCTGCTTGTGGCTTCTATTTTC 3420
Db 3361 AGCGGCCGAGAGCATCTCTCTCTTACAAAATCTGCAATCTGCTTGTGGCTTCTATTTTC 3420
Qy 3421 CAAATAAAGAAAGTCTAAATTTGAGCACCGCAAGGTGCACACCAAAAAAATCTGCTTTCCGTA 3480
Db 3421 CAAATAAAGAAAGTCTAAATTTGAGCACCGCAAGGTGCACACCAAAAAAATCTGCTTTCCGTA 3480
Qy 3481 CCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGATGCGCTCTCCAGGAGGAGCTTCC 3540
Db 3481 CCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGATGCGCTCTCCAGGAGGAGCTTCC 3540
Qy 3541 TGCAGTTGTTCAACTTGAGACCCAAATCTCACCTGAAACGGGGAAGAGCCCTGTGAGAT 3600
Db 3541 TGCAGTTGTTCAACTTGAGACCCAAATCTCACCTGAAACGGGGAAGAGCCCTGTGAGAT 3600
Qy 3601 GCATCCCTCAGCTCGATCCGTTTCCACCCTTCCAGGCTTTGGCAGCTGGCTTACCAAGGAA 3660
Db 3601 GCATCCCTCAGCTCGATCCGTTTCCACCCTTCCAGGCTTTGGCAGCTGGCTTACCAAGGAA 3660
Qy 3661 AAGTTGCCATTTTCCAAAGATGAGGAAATCGGGCAAGAGGAGCAACGCAACGACG 3720
Db 3661 AAGTTGCCATTTTCCAAAGATGAGGAAATCGGGCAAGAGGAGCAACGCAACGACG 3720
Qy 3721 ATTGAGTTCCGAGAAAGGAGCTTTGAGAAACAAATAAGGGCAGTTGTGAGGCTCTCGC 3780
Db 3721 ATTGAGTTCCGAGAAAGGAGCTTTGAGAAACAAATAAGGGCAGTTGTGAGGCTCTCGC 3780
Qy 3781 AAGAGAAAGAGAGTGCAAACACTCCACGGCGAAGCGCCCTCCGTTGAGCGCGGATCCCA 3840
Db 3781 AAGAGAAAGAGAGTGCAAACACTCCACGGCGAAGCGCCCTCCGTTGAGCGCGGATCCCA 3840
Qy 3841 AGTTACCCAGTAGCAGGAGAGCCCACTCACTGCTCCGAGTGCAGCAAGCTTTTCAGAA 3900
Db 3841 AGTTACCCAGTAGCAGGAGAGCCCACTCACTGCTCCGAGTGCAGCAAGCTTTTCAGAA 3900
Qy 3901 CCTACCAACAGCTGTTTGTGCACTCCAGGCTTCAAGAGGACCGGAGGCGCGCGG 3960
Db 3901 CCTACCAACAGCTGTTTGTGCACTCCAGGCTTCAAGAGGACCGGAGGCGCGCGG 3960

QY 3961 AGTCGCCACCATGTCTGTGGACGGGAGCGACGCCGGGACGTGTTCTCTCGACCTCGCG 4020
DB |||||
DB 3961 AGTCGCCACCATGTCTGTGGACGGGAGCGACGCCGGGACGTGTTCTCTCGACCTCGCG 4020
QY 4021 CCCCTCTGGATGAAATGTGAGCCGTGGATCTGAGGGGAAAGTGTTCTGAAAGCGGATCTG 4080
DB |||||
DB 4021 CCCCTCTGGATGAAATGTGAGCCGTGGATCTGAGGGGAAAGTGTTCTGAAAGCGGATCTG 4080
QY 4081 AGGATGGGCTCCCGAAGAAATCCATCTGGGTAAGCTGCGCTCTCGTCCCGTCTGTGT 4140
DB |||||
DB 4081 AGGATGGGCTCCCGAAGAAATCCATCTGGGTAAGCTGCGCTCTCGTCCCGTCTGTGT 4140
QY 4141 TCCGCTGTGTCTGTCTGTCTCCCGTCTCCCGTCTCTATTTCCCATCTCTCAGACAAACGC 4200
DB |||||
DB 4141 TCCGCTGTGTCTGTCTGTCTCCCGTCTCCCGTCTCTATTTCCCATCTCTCAGACAAACGC 4200
QY 4201 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTGGTATCACTCT 4260
DB |||||
DB 4201 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTGGTATCACTCT 4260
QY 4261 GTGTAAATCAATTAACCTCTCAGGGCTTAATTTTCTCATTTCTGTATAACAGGGTTGA 4320
DB |||||
DB 4261 GTGTAAATCAATTAACCTCTCAGGGCTTAATTTTCTCATTTCTGTATAACAGGGTTGA 4320
QY 4321 GTTAAGAGTCTCTGTGTTCTGAAATATATATATATTTTAAACGTGTATCGTTTGC 4380
DB |||||
DB 4321 GTTAAGAGTCTCTGTGTTCTGAAATATATATATATTTTAAACGTGTATCGTTTGC 4380
QY 4381 TCACAAAACACACTTTTAAAAAAAATAACTTGTGCATCCAGCCCAAAATGCATCTCTCT 4440
DB |||||
DB 4381 TCACAAAACACACTTTTAAAAAAAATAACTTGTGCATCCAGCCCAAAATGCATCTCTCT 4440
QY 4441 TAACTGGGGCGATTTTGTTCCTCAATCAGTATCTGGCAATGTCTGGAGGCAATTTGGTTGT 4500
DB |||||
DB 4441 TAACTGGGGCGATTTTGTTCCTCAATCAGTATCTGGCAATGTCTGGAGGCAATTTGGTTGT 4500
QY 4501 CATACTGTGTGTGGGTGCTCTCTGTCATCCAGTGGGAGGAGCCAGGAGCACTGCT 4560
DB |||||
DB 4501 CATACTGTGTGTGGGTGCTCTCTGTCATCCAGTGGGAGGAGCCAGGAGCACTGCT 4560
QY 4561 CAGCATGTGTACAGTGCACAGGACAGCCCATCATCAAGAAATTAATCTGCTCCCAATGTC 4620
DB |||||
DB 4561 CAGCATGTGTACAGTGCACAGGACAGCCCATCATCAAGAAATTAATCTGCTCCCAATGTC 4620
QY 4621 AATAGTTGAGCAATTTAGGAGACCTTAGCCTTCACTTAAGTTTCTGCGGTTCTGATCT 4680
DB |||||
DB 4621 AATAGTTGAGCAATTTAGGAGACCTTAGCCTTCACTTAAGTTTCTGCGGTTCTGATCT 4680
QY 4681 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAGGTACTGGGAGTGTATCAACTAGAGCCAG 4740
DB |||||
DB 4681 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAGGTACTGGGAGTGTATCAACTAGAGCCAG 4740
QY 4741 GAATATTAATTTGGCAGCGTTTGGTGTCTCCAAACCTTCTCTCTGTCTGGCAAG 4800
DB |||||
DB 4741 GAATATTAATTTGGCAGCGTTTGGTGTCTCCAAACCTTCTCTCTGTCTGGCAAG 4800
QY 4801 CTAGTATCCATTTATAGGTACCTCAGGAAACCAATATGATTTGTCTATAAATAAAGGAAT 4860
DB |||||
DB 4801 CTAGTATCCATTTATAGGTACCTCAGGAAACCAATATGATTTGTCTATAAATAAAGGAAT 4860
QY 4861 GTGAGCACACTGAAGACATTTTAAAGAGGCTCATTTGCTCAGCAGAAATTTTCAGTGTAC 4920
DB |||||
DB 4861 GTGAGCACACTGAAGACATTTTAAAGAGGCTCATTTGCTCAGCAGAAATTTTCAGTGTAC 4920
QY 4921 TAGTGGCAATTTATAGAAAGAGAGGTGATCACTGAAGGCATGCTCACATAAATATTTCTGA 4980
DB |||||
DB 4921 TAGTGGCAATTTATAGAAAGAGAGGTGATCACTGAAGGCATGCTCACATAAATATTTCTGA 4980
QY 4981 GCCCTGGTGGGGTTATCTAGGGCAAGGAATTCACCTGTGTGTGGAGTTGGCCCATCC 5040
DB |||||
DB 4981 GCCCTGGTGGGGTTATCTAGGGCAAGGAATTCACCTGTGTGTGGAGTTGGCCCATCC 5040
QY 5041 TCACTGTAGCCAGAGCTTCTCTATCAGAGTTTAGTATTTTGTGAAATAGAGGATCTTG 5100

DB 5041 TCACGTAGCCAGAGCTTCTCTATCAGATTAGTATTTTGTGTAATAGAGGATCTTG 5100
QY 5101 CTGCTTAAAAAGAGTTGAAAGACCTGATGGGAGCGCGTAATTAAGCAAGCGAATGATGG 5160
DB |||||
DB 5101 CTGCTTAAAAAGAGTTGAAAGACCTGATGGGAGCGCGTAATTAAGCAAGCGAATGATGG 5160
QY 5161 GAAATGATGATCGGCTTTAGGGAAGCATCTGTCAAAGTGGTCTTGGTTAAACAAAGTGGC 5220
DB |||||
DB 5161 GAAATGATGATCGGCTTTAGGGAAGCATCTGTCAAAGTGGTCTTGGTTAAACAAAGTGGC 5220
QY 5221 TCCTCTCTCAGTGTCACTTGAATTTGCTGCTTGAATTTCTCGGAAAACTGGGTGTATGAG 5280
DB |||||
DB 5221 TCCTCTCTCAGTGTCACTTGAATTTGCTGCTTGAATTTCTCGGAAAACTGGGTGTATGAG 5280
QY 5281 ACCACAGATGAATTTGGCCACACAGATTGATTGGACTCTTCTCTCACTGCTCTTCAGCCA 5340
DB |||||
DB 5281 ACCACAGATGAATTTGGCCACACAGATTGATTGGACTCTTCTCTCACTGCTCTTCAGCCA 5340
QY 5341 GTGCCAGTTCCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCAAAATC 5400
DB |||||
DB 5341 GTGCCAGTTCCCTTTTCTGATCATGTGATTGACGTGAGAACTGTAGTCTGTATATCAAAATC 5400
QY 5401 TTTAGAAATGTTTTTGGAGTTTCTGGGACACAGGAAACCCAGCACTTAGCATACTACAAT 5460
DB |||||
DB 5401 TTTAGAAATGTTTTTGGAGTTTCTGGGACACAGGAAACCCAGCACTTAGCATACTACAAT 5460
QY 5461 CTAATGTCTTAATGGCATCAATAAAGAGGCTTTAAAAACAGAGTCTCAGTCTAGTAAAGTG 5520
DB |||||
DB 5461 CTAATGTCTTAATGGCATCAATAAAGAGGCTTTAAAAACAGAGTCTCAGTCTAGTAAAGTG 5520
QY 5521 GTTTCGTGTAGTCCGGTACTGTTGACAGGGGCCCTGTGAGTGGCCAGTTCCTCTGAAAG 5580
DB |||||
DB 5521 GTTTCGTGTAGTCCGGTACTGTTGACAGGGGCCCTGTGAGTGGCCAGTTCCTCTGAAAG 5580
QY 5581 AAATGAAAGGCGAGTTACCGGTAGTGTGTGGAACCATGGGCTAGATCATCAGGCAG 5640
DB |||||
DB 5581 AAATGAAAGGCGAGTTACCGGTAGTGTGTGGAACCATGGGCTAGATCATCAGGCAG 5640
QY 5641 GACGAATGCTGTGGTGTGGGAGCAACCCAGCTTGGGCTTGGTCTGATCTGCTTCTAC 5700
DB |||||
DB 5641 GACGAATGCTGTGGTGTGGGAGCAACCCAGCTTGGGCTTGGTCTGATCTGCTTCTAC 5700
QY 5701 ACTGGTGTGTGTGACCAATTAAGTGTGCTTAACTTTCTTGTCTACTATTTCCCTG 5760
DB |||||
DB 5701 ACTGGTGTGTGTGACCAATTAAGTGTGCTTAACTTTCTTGTCTACTATTTCCCTG 5760
QY 5761 TTTGCAAAATGGTTCAITGACCCCTGTCTCCACCTCCCAAGGACAAATTTCAACAGCCTA 5820
DB |||||
DB 5761 TTTGCAAAATGGTTCAITGACCCCTGTCTCCACCTCCCAAGGACAAATTTCAACAGCCTA 5820
QY 5821 TTTGTAAGAAAGATCAAGTCTTTTAAATAATACTGTAAAGTCAAGGAGTGTGATGGA 5880
DB |||||
DB 5821 TTTGTAAGAAAGATCAAGTCTTTTAAATAATACTGTAAAGTCAAGGAGTGTGATGGA 5880
QY 5881 AAGACGGAACCCAGGTAGTGTGAAATGTCTGCTTTGTTCTTAAAGAAAGGCATT 5940
DB |||||
DB 5881 AAGACGGAACCCAGGTAGTGTGAAATGTCTGCTTTGTTCTTAAAGAAAGGCATT 5940
QY 5941 TCATAGCTTTTGGATATGACGCAACATACCAATAATCTGACACATAGTTGGGAGTGG 6000
DB |||||
DB 5941 TCATAGCTTTTGGATATGACGCAACATACCAATAATCTGACACATAGTTGGGAGTGG 6000
QY 6001 AAATGGCAACACGCCCGGATTATAAACCCAGGTAGTTGGGTATGATTTGTAAGAAAAA 6060
DB |||||
DB 6001 AAATGGCAACACGCCCGGATTATAAACCCAGGTAGTTGGGTATGATTTGTAAGAAAAA 6060
QY 6061 AGCTGGCCATTTCTGATTTGGGGAATGATTTTCTTAACTTATATTTCTTAGTAGTCT 6120
DB |||||
DB 6061 AGCTGGCCATTTCTGATTTGGGGAATGATTTTCTTAACTTATATTTCTTAGTAGTCT 6120
QY 6121 AGATTTATCATATTTGATCTATCATCTGCTTTTAAAGCTTAAAGAAATCAAGTAAAT 6180
DB |||||

Db 6121 AGATTTATCATATGTACTATCATCTCGGCTTTTTTAAAGACTTAAAGAGATCAAGTAAAT 6180
Qy TTTTTCCTTCCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTACAGGTGGAT 6240
Db TTTTTCCTTCCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTACAGGTGGAT 6240
Qy AGTGATATGATCTACAGTGAGGGGACATTTATTTTAAACCTTAAACATTCATGTGTTTGG 6300
Db AGTGATATGATCTACAGTGAGGGGACATTTATTTTAAACCTTAAACATTCATGTGTTTGG 6300
Qy GGTGTGTATTTTAAACGGACACTCTGATGTCTTTTGGAGGGCTGGTGTGTGTTGAA 6360
Db GGTGTGTATTTTAAACGGACACTCTGATGTCTTTTGGAGGGCTGGTGTGTGTTGAA 6360
Qy GTTCTGTCTCTTCCTTCAGGTGACCTCTAACTTCTCTGTATGTCAGTGCAGACACATTCCT 6420
Db GTTCTGTCTCTTCCTTCAGGTGACCTCTAACTTCTCTGTATGTCAGTGCAGACACATTCCT 6420
Qy ATTGTCTCTGAGAAACTTAAAGCCAAACACCTGTCTATCTGGGACAGGTTTTCAITTTGTAG 6480
Db ATTGTCTCTGAGAAACTTAAAGCCAAACACCTGTCTATCTGGGACAGGTTTTCAITTTGTAG 6480
Qy ATCTCTTTGCCCCACATGAGTGTGTTGGACAATACAGCTGTCTTCCAAAACCTTTGCTA 6540
Db ATCTCTTTGCCCCACATGAGTGTGTTGGACAATACAGCTGTCTTCCAAAACCTTTGCTA 6540
Qy AATTTTGACAGACTTTCTTAGTGTCTTGGCCAAATGCGAGACTTTCTTTCTGTGGAAGAT 6600
Db AATTTTGACAGACTTTCTTAGTGTCTTGGCCAAATGCGAGACTTTCTTTCTGTGGAAGAT 6600
Qy TAAGTTGTCTGTCTGCCCTCTAGTGTCTAGTGTGTTTAACTCTTAAACGGCTTAT 6660
Db TAAGTTGTCTGTCTGCCCTCTAGTGTCTAGTGTGTTTAACTCTTAAACGGCTTAT 6660
Qy TTTTCCCTCTGGTGTGGAGTTGACGGTTGTAAATGTGCTCATTTTCTTAAATTTATTC 6720
Db TTTTCCCTCTGGTGTGGAGTTGACGGTTGTAAATGTGCTCATTTTCTTAAATTTATTC 6720
Qy TGAAGAAGATAATTTTTCGCCGAGTATGTATGTCACCTTCACTTGTGCGAGATCCTGCC 6780
Db TGAAGAAGATAATTTTTCGCCGAGTATGTATGTCACCTTCACTTGTGCGAGATCCTGCC 6780
Qy TGCTCAGAGACACTGAGAACCGGAAGCTGCGCGGCAATTCAGTCTATGAATGATCTTT 6840
Db TGCTCAGAGACACTGAGAACCGGAAGCTGCGCGGCAATTCAGTCTATGAATGATCTTT 6840
Qy CTTGTGATTAAGCAACGAGAACCTGAATGTTTAAATAGTGTACTCTGTGTACCCAGAA 6900
Db CTTGTGATTAAGCAACGAGAACCTGAATGTTTAAATAGTGTACTCTGTGTACCCAGAA 6900
Qy AAAAAACAAACAAATCATGTTTATACACTCTAABACTTCAACACCTTCAACAGCATT 6960
Db AAAAAACAAACAAATCATGTTTATACACTCTAABACTTCAACACCTTCAACAGCATT 6960
Qy TGGTGTGTCTAGCGGTTTTGTTCTAACCCGATGTTATATAAGAAATTTTTTCATGCT 7020
Db TGGTGTGTCTAGCGGTTTTGTTCTAACCCGATGTTATATAAGAAATTTTTTCATGCT 7020
Qy TTCCAAAATGTTATGTCAAGATAATTTAAGTCAGCTGCTTATTCAGGTACTTTCAGC 7080
Db TTCCAAAATGTTATGTCAAGATAATTTAAGTCAGCTGCTTATTCAGGTACTTTCAGC 7080
Qy TACCTTCTTATATAATTTTTTGTGTTTCTTTTAAAGATAAAATGATCATGAGGAAA 7140
Db TACCTTCTTATATAATTTTTTGTGTTTCTTTTAAAGATAAAATGATCATGAGGAAA 7140
Qy ATAAAAATCTTACATCTTCAAGAGAGTGTAGTTATTTGGAAAGTTTTTCGGTTCAAAT 7200
Db ATAAAAATCTTACATCTTCAAGAGAGTGTAGTTATTTGGAAAGTTTTTCGGTTCAAAT 7200
Qy TATTACCTCAATTTCACTCAGAACGCAATACAGGTAAAGAACTTTTATTTTTTAAACA 7260
Db TATTACCTCAATTTCACTCAGAACGCAATACAGGTAAAGAACTTTTATTTTTTAAACA 7260

Qy TGCATTAGTTAAATATGTAGTTATCTAAATTTTTTTTGTGTTCTGTTTCTGACTCTGC 7320
Db TGCATTAGTTAAATATGTAGTTATCTAAATTTTTTTTGTGTTCTGTTTCTGACTCTGC 7320
Qy CAGATCCTTTGGACTAGCTTTAAGGATAAATATGTAGCATGTGTAGTGCAGTGGTATTTTT 7380
Db CAGATCCTTTGGACTAGCTTTAAGGATAAATATGTAGCATGTGTAGTGCAGTGGTATTTTT 7380
Qy ATTCCTTTAGTGCCATTGTAACCTTGAGCCATTGTCTTATTTTCAGTTCAATTTCTTTCT 7440
Db ATTCCTTTAGTGCCATTGTAACCTTGAGCCATTGTCTTATTTTCAGTTCAATTTCTTTCT 7440
Qy TTTCTTTTGTGTTTTTGGACGAGTCTTGTCTGTCACTCGGCTGGAGTGCAGTGGTG 7500
Db TTTCTTTTGTGTTTTTGGACGAGTCTTGTCTGTCACTCGGCTGGAGTGCAGTGGTG 7500
Qy CAATTTTCGCTCACTGACGCTCCACCTCCCTGGTTCAAGCAATACTCTGCTCAGCT 7560
Db CAATTTTCGCTCACTGACGCTCCACCTCCCTGGTTCAAGCAATACTCTGCTCAGCT 7560
Qy CCCCAGTAGTTGGGATTAACAGGTACCTGCGCACCAACCCGGCTAAATTTCTGTATTTTTAG 7620
Db CCCCAGTAGTTGGGATTAACAGGTACCTGCGCACCAACCCGGCTAAATTTCTGTATTTTTAG 7620
Qy TAGAGATGGGTTTCAACCATGCTGGCCAGGCTGTTTTCGAACTCTCGACCTCAAGTGATC 7680
Db TAGAGATGGGTTTCAACCATGCTGGCCAGGCTGTTTTCGAACTCTCGACCTCAAGTGATC 7680
Qy CCGTCACTTGGCTCCCATAGTGTGGCTCCCATAGTGTGGGATTAACAGCGTGAGC 7740
Db CCGTCACTTGGCTCCCATAGTGTGGCTCCCATAGTGTGGGATTAACAGCGTGAGC 7740
Qy CACCGGCGCGGACAAAGTTCAATTTGTAGTTTATGACTGTCTATGCTGACTCTTATC 7800
Db CACCGGCGCGGACAAAGTTCAATTTGTAGTTTATGACTGTCTATGCTGACTCTTATC 7800
Qy TTATTTAAAGCTTACAGTATTTTAAATGCTGCATCTTATGCTTATGATGAGAATGAA 7860
Db TTATTTAAAGCTTACAGTATTTTAAATGCTGCATCTTATGCTTATGATGAGAATGAA 7860
Qy ATGAGATCTATTTAGTAGTCTTGAGATTGTGAAGGAGCTATGACATCATGATGAGGA 7920
Db ATGAGATCTATTTAGTAGTCTTGAGATTGTGAAGGAGCTATGACATCATGATGAGGA 7920
Qy GCGTGGTAGATTGAAATTTTCATCTTCCACTTACTATCTGTGACCCCTTTGGGCAAGT 7980
Db GCGTGGTAGATTGAAATTTTCATCTTCCACTTACTATCTGTGACCCCTTTGGGCAAGT 7980
Qy TATTTTAAACCTTTTGTGCTTTTGTAGTTTCTTTTGTGTAAGGATAGATAATCATATTTT 8040
Db TATTTTAAACCTTTTGTGCTTTTGTAGTTTCTTTTGTGTAAGGATAGATAATCATATTTT 8040
Qy CTAGGCTGTTTGAAGATTAAATAGTTAGAGTGTGCTGTTTAAATTTTTTCTTATGAA 8100
Db CTAGGCTGTTTGAAGATTAAATAGTTAGAGTGTGCTGTTTAAATTTTTTCTTATGAA 8100
Qy GATAGGCATTCAATATTTCAATATTTTATTACAGTAAGATGATAAAGAACTGATGAGAA 8160
Db GATAGGCATTCAATATTTCAATATTTTATTACAGTAAGATGATAAAGAACTGATGAGAA 8160
Qy ATCTCTATGTAGTAGATCGAGAAAGCAAAAGGAGGAAAGGCTGTTTTCTTAAATAA 8220
Db ATCTCTATGTAGTAGATCGAGAAAGCAAAAGGAGGAAAGGCTGTTTTCTTAAATAA 8220
Qy ATAGATAATTTGATCTATTTTTCAGTGTCTTTTCACTCTTCTATAAATAAGTGCCATTTCT 8280
Db ATAGATAATTTGATCTATTTTTCAGTGTCTTTTCACTCTTCTATAAATAAGTGCCATTTCT 8280
Qy GCGTTAGGTGAAACCAATACAAATGTGAATTTTGTGAATATGCTGCAGCCAGAGACA 8340
Db GCGTTAGGTGAAACCAATACAAATGTGAATTTTGTGAATATGCTGCAGCCAGAGACA 8340

QY	2281	THAAAGTGAGCAATGAAATAGGTAGTACCTTCTCTGCTGGTGGAAAAAGACCGTATGACTC	2340	QY	3361	ACGGCGCGAGAGCATCTCTCTCTTACAAAACTCGCATGTTGTGGCTTCTCTATTTC	3420
Db	9808	TAAGAGTGAAGCAATGAAATAGGTAGTACCTTCTCTGCTGGTGGAAAAAGACCGTATGACTC	9867	Db	10888	ACGGCGCGAGAGCATCTCTCTCTTACAAAACTCGCATGTTGTGGCTTCTCTATTTC	10947
QY	2341	ACCCACACGAGCTTCTCTCTGCTCTGAGTGAGCTAACCGTTTCTGTTTTTTTCTCTCT	2400	QY	3421	CAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAATCTGTTTCGGTA	3480
Db	9868	ACCCACACGAGCTTCTCTCTGCTCTGAGTGAGCTAACCGTTTCTGTTTTTTTCTCTCT	9927	Db	10948	CAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAATCTGTTTCGGTA	11007
QY	2401	AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA	2460	QY	3481	CCAGAGCGCGCAGACAGACTCTCCAAAGAGGAATCCGCTCTCTGAGGAGGAGCTTCC	3540
Db	9928	AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACA	9987	Db	11008	CCAGAGCGCGCAGACAGACTCTCCAAAGAGGAATCCGCTCTCTGAGGAGGAGCTTCC	11067
QY	2461	CTCAATTCAGTGAAGCAAGAACTAATGCTTTGTGCTGATTCATATTTGAAATCGAGGC	2520	QY	3541	TGCAGTTGTTCAACTTGAGACCAAAATCTCACCTGAAACCGGGGAAGAGCTGTCAAGT	3600
Db	9988	CTCAATTCAGTGAAGCAAGAACTAATGCTTTGTGCTGATTCATATTTGAAATCGAGGC	10047	Db	11068	TGCAGTTGTTCAACTTGAGACCAAAATCTCACCTGAAACCGGGGAAGAGCTGTCAAGT	11127
QY	2521	ATTGGGAACCTGTATGCTTGTGTTGTGGAAGAACAGTGCACCATCACTGAGCTTCC	2580	QY	3601	GCATCCCTCAGTCCGTTCCACACTTTCAGGCTTGGCAGCTGGCTACCAAGGAA	3660
Db	10048	ATTGGGAACCTGTATGCTTGTGTTGTGGAAGAACAGTGCACCATCACTGAGCTTCC	10107	Db	11128	GCATCCCTCAGTCCGTTCCACACTTTCAGGCTTGGCAGCTGGCTACCAAGGAA	11187
QY	2581	TAAAGTTTGAAGAGTTAGGAGTATACACTTTCTTTTGAATCTTTTATAATAATATT	2640	QY	3661	AAGTTGCCATTTCGCAAGAGTGAAGGAATCGGGCAAGAGGAGGAGCACCGAACGAG	3720
Db	10108	TAAAGTTTGAAGAGTTAGGAGTATACACTTTCTTTTGAATCTTTTATAATAATATT	10167	Db	11188	AAGTTGCCATTTCGCAAGAGTGAAGGAATCGGGCAAGAGGAGGAGCACCGAACGAG	11247
QY	2641	TGCTCTGGTTTGGAAACCCAGGGCTGTAGAGGGTGAGTGACAAAGTCTTACAAGTGCC	2700	QY	3721	ATTTCAGTTCCGGAAGAGGAGCTTGGAGAAACAAATAAGGCGAGTTGTGAGGCTCTCGC	3780
Db	10168	TGCTCTGGTTTGGAAACCCAGGGCTGTAGAGGGTGAGTGACAAAGTCTTACAAGTGCC	10227	Db	11248	ATTTCAGTTCCGGAAGAGGAGCTTGGAGAAACAAATAAGGCGAGTTGTGAGGCTCTCGC	11307
QY	2701	CTTATTCCAACTCCAGAAATTCGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA	2760	QY	3781	AAGGAAAGAGAGTGCACCACTCCACGGCGAAGCGCCCTCCGTGGACGGGATCCCA	3840
Db	10228	CTTATTCCAACTCCAGAAATTCGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGA	10287	Db	11308	AAGGAAAGAGAGTGCACCACTCCACGGCGAAGCGCCCTCCGTGGACGGGATCCCA	11367
QY	2761	CAGGAAACATGCAACTCAATCCCTTAAATGTACATGAGTGGGCCAGAGTGTATGGCA	2820	QY	3841	AGTTACCCAGTAGCAAGAGAGGCCACTACTGTCTCGAGTGGCGGCAAGCTTTTCAGAA	3900
Db	10288	CAGGAAACATGCAACTCAATCCCTTAAATGTACATGAGTGGGCCAGAGTGTATGGCA	10347	Db	11368	AGTTACCCAGTAGCAAGAGAGGCCACTACTGTCTCGAGTGGCGGCAAGCTTTTCAGAA	11427
QY	2821	GCTCTCTGGAGTCCGATGAGATGAGGATGCCCTGTGTCAATGAAGGACCGCTGTG	2880	QY	3901	CCTACCAACAGCTGTGTGTGCACTCCAGGTTCCACAGAGGACCGGAGGCGCGCGG	3960
Db	10348	GCTCTCTGGAGTCCGATGAGATGAGGATGCCCTGTGTCAATGAAGGACCGCTGTG	10407	Db	11428	CCTACCAACAGCTGTGTGTGCACTCCAGGTTCCACAGAGGACCGGAGGCGCGCGG	11487
QY	2881	TTCCATTCGAGCTACACAGAAAAAATGTCTCAAAATCGAGGGTATATGCCCTTGG	2940	QY	3961	AGTGCCTCCACCATGTCTGTGACGGGAGCGCGGACGTGTCTCTGACCTCGCG	4020
Db	10408	TTCCATTCGAGCTACACAGAAAAAATGTCTCAAAATCGAGGGTATATGCCCTTGG	10467	Db	11488	AGTGCCTCCACCATGTCTGTGACGGGAGCGCGGACGTGTCTCTGACCTCGCG	11547
QY	2941	ATTGCATGTTCTGCAGCCAGACTTTCACACATTCAGAGACCTTAAATAACATGCTTAA	3000	QY	4021	CCCTCTCGATGAAAAATGGAGCGGTGATCGAGGGGAAAGTGGTCTGAAAGCGATCTG	4080
Db	10468	ATTGCATGTTCTGCAGCCAGACTTTCACACATTCAGAGACCTTAAATAACATGCTTAA	10527	Db	11548	CCCTCTCGATGAAAAATGGAGCGGTGATCGAGGGGAAAGTGGTCTGAAAGCGATCTG	11607
QY	3001	TGCAACACGGCTTACCTCTGTGAACCGAGATTCCTCGGGTGAAGCAGAGTATCTCA	3060	QY	4081	AGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAAGTCCCTGTCTCGTCCCGTGTGT	4140
Db	10528	TGCAACACGGCTTACCTCTGTGAACCGAGATTCCTCGGGTGAAGCAGAGTATCTCA	10587	Db	11608	AGGATGGGCTTCCCGAAGGAATCCATCTGGGTAAAGTCCCTGTCTCGTCCCGTGTGT	11667
QY	3061	GTCCGCTTGATAAAAGTCAAGTGCAGAACAGAACCTCCCAAGGAAAGAAATTCAGAGGAA	3120	QY	4141	TCCGCTGTGTCTGTCTCTCCCGTCTCCCGCTCTCTATTTCCCATCTCTCCAGAACGC	4200
Db	10588	GTCCGCTTGATAAAAGTCAAGTGCAGAACAGAACCTCCCAAGGAAAGAAATTCAGAGGAA	10647	Db	11668	TCCGCTGTGTCTGTCTCTCCCGTCTCCCGCTCTCTATTTCCCATCTCTCCAGAACGC	11727
QY	3121	ATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCCCTTTTGTGATGATCC	3180	QY	4201	TGGCAGGAAATGGGGTTTGGAGAGCCAGTCAAGTCCAGGCTCTTTTGGTATCACTCT	4260
Db	10648	ATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCCCTTTTGTGATGATCC	10707	Db	11728	TGGCAGGAAATGGGGTTTGGAGAGCCAGTCAAGTCCAGGCTCTTTTGGTATCACTCT	11787
QY	3181	ACATGAGAAACACAAAGATTTCTTCACTTACGGGTGTAACATGTGCGGAAGAAATTC	3240	QY	4261	GTGTAAGTCAATTAACCTCTCAGGCTTAAATTTCTCATTTCTGTAATAACAGGTTGA	4320
Db	10708	ACATGAGAAACACAAAGATTTCTTCACTTACGGGTGTAACATGTGCGGAAGAAATTC	10767	Db	11788	GTGTAAGTCAATTAACCTCTCAGGCTTAAATTTCTCATTTCTGTAATAACAGGTTGA	11847
QY	3241	AGGAGCTTGGTTTCTTAAAAATCACATGCGGACACATATGCGCAATCGGGGCGAGAA	3300	QY	4321	GTTAAGAGTCTCTCTGTTTCTGAAATATATATATATTTTAAACGTGTATCGTTTTGC	4380
Db	10768	AGGAGCTTGGTTTCTTAAAAATCACATGCGGACACATATGCGCAATCGGGGCGAGAA	10827	Db	11848	GTTAAGAGTCTCTCTGTTTCTGAAATATATATATATTTTAAACGTGTATCGTTTTGC	11907
QY	3301	GCAAACTGCAAGGCTTGGAGTGTCCAGCAACGATCAACGAGGTCTGTCAGGTTGC	3360	QY	4381	TCACAAAAACACTTTTAAAAAATAAATTAACCTTGTCATCCAGCCCAAAATGCACTGTTCT	4440
Db	10828	GCAAACTGCAAGGCTTGGAGTGTCCAGCAACGATCAACGAGGTCTGTCAGGTTGC	10887	Db	11908	TCACAAAAACACTTTTAAAAAATAAATTAACCTTGTCATCCAGCCCAAAATGCACTGTTCT	11967
				QY	4441	TAACTGGGGCGATTTTGTTCCTCCCAATCAGTATCTGCGCAATGTCTGGAGGCAATTTTGGTTCT	4500

Db	11968	TAAC	TGGGGCGA	TTTTGT	CCCAAT	CAGTAT	TGGCAA	TGTCT	TGAGAGCA	TTTGTGGTTGT	12027	
Qy	4501	CATA	CTGTGTGT	TGGGTGT	GCCTGT	GTGGCAT	CCAGTGGCGC	GAGAGCGC	CAGGACAC	TGCTGCT	4560	
Db	12028	CATA	CTGTGTGT	TGGGTGT	GCCTGT	GTGGCAT	CCAGTGGCGC	GAGAGCGC	CAGGACAC	TGCTGCT	12087	
Qy	4561	CAG	CATGGTA	CAGTGC	CACAGGA	CAGCCCCAT	CAATCA	AAAGAA	TTATCT	TGTCGCCAAATGTC	4620	
Db	12088	CAG	CATGGTA	CAGTGC	CACAGGA	CAGCCCCAT	CAATCA	AAAGAA	TTATCT	TGTCGCCAAATGTC	12147	
Qy	4621	AAT	AGTTTT	TGAGCA	TTGAGAGAC	CCCTAG	CCCTTCA	TTTAAG	TTTTCT	TGCGGTTCTTGATCT	4680	
Db	12148	AAT	AGTTTT	TGAGCA	TTGAGAGAC	CCCTAG	CCCTTCA	TTTAAG	TTTTCT	TGCGGTTCTTGATCT	12207	
Qy	4681	TTTT	CTGTAG	TGAATTT	CTAGTGGCC	ATAAAG	GTACTGGG	AG---	TGATCA	CAACTAGAGC	4737	
Db	12208	TTTT	CTGTAG	TGAATTT	CTAGTGGCC	ATAAAG	GTACTGGG	AG---	TGATCA	CAACTAGAGC	12267	
Qy	4738	CAG	GAATAT	TATTT	TGGG	CAGCCG	TTTGGT	GTGTCA	AAAA	CCCTGTCTTCTGTCTGGC	4797	
Db	12268	CAG	GAATAT	TATTT	TGGG	CAGCCG	TTTGGT	GTGTCA	AAAA	CCCTGTCTTCTGTCTGGC	12327	
Qy	4798	AAG	CTAGTAT	CCATTT	ATATAG	GTACCT	CAGGA	CCCCAAT	TGATTT	GTGCATAAATAC	4857	
Db	12328	AAG	CTAGTAT	CCATTT	ATATAG	GTACCT	CAGGA	CCCCAAT	TGATTT	GTGCATAAATAC	12387	
Qy	4858	AAT	GTGAG	CAC	CTG	GAAG	ACATTTT	TAA	AGGCTCAT	TTGCTCAGCAGAA	4917	
Db	12388	AA	TGTGAG	CAC	CTG	GAAG	ACATTTT	TAA	AGGCTCAT	TTGCTCAGCAGAA	12447	
Qy	4918	TAC	TAGTGG	CA	TTAT	TAGAA	AGAG	AGTGAT	CAC	TGAAGGCAT	5097	
Db	12448	TAC	TAGTGG	CA	TTAT	TAGAA	AGAG	AGTGAT	CAC	TGAAGGCAT	12507	
Qy	4978	TGA	CCCTGT	GTGG	CGTTAT	CTAG	GGCA	AAAGAT	TTCC	ACCTGTGTTTGAGTTGGGCCA	5037	
Db	12508	TGA	CCCTGT	GTGG	CGTTAT	CTAG	GGCA	AAAGAT	TTCC	ACCTGTGTTTGAGTTGGGCCA	12567	
Qy	5038	TC	CTACTGT	PAG	CCAG	AGCTTCT	CTCAT	CAGAG	TTTGTAG	TATTTGTTTGAATAC	5097	
Db	12568	TC	CTACTGT	PAG	CCAG	AGCTTCT	CTCAT	CAGAG	TTTGTAG	TATTTGTTTGAATAC	12627	
Qy	5098	TT	CTGCTT	AAAA	CAGTT	TGAA	AGAC	CCCTG	ATG	TGGCGAGCCGTA	5157	
Db	12628	TT	CTGCTT	AAAA	CAGTT	TGAA	AGAC	CCCTG	ATG	TGGCGAGCCGTA	12687	
Qy	5158	TGG	GAACAT	GAAT	TCGGTCT	TTAG	GAAG	CAATCT	GTG	CAAA	GTGGTCTTGGTTAA	5217
Db	12688	TGG	GAACAT	GAAT	TCGGTCT	TTAG	GAAG	CAATCT	GTG	CAAA	GTGGTCTTGGTTAA	12747
Qy	5218	GC	CTCTCT	CTCT	CAG	GTGC	CACTT	GATGTG	CTCT	GAATTTCTT	5277	
Db	12748	GC	CTCTCT	CTCT	CAG	GTGC	CACTT	GATGTG	CTCT	GAATTTCTT	12807	
Qy	5278	GAG	ACCCAG	AT	GAAATTT	TGCC	CAC	AGATTG	AT	TGGA	5337	
Db	12808	GAG	ACCCAG	AT	GAAATTT	TGCC	CAC	AGATTG	AT	TGGA	12867	
Qy	5338	CA	GTGCG	AGTT	CTCTTT	CTG	ATCAT	GTGAT	TGA	CGTG	5397	
Db	12868	CA	GTGCG	AGTT	CTCTTT	CTG	ATCAT	GTGAT	TGA	CGTG	12927	
Qy	5398	AT	CTTT	TAG	AATG	TTTT	TG	AGTTT	CT	TGG	5457	
Db	12928	AT	CTTT	TAG	AATG	TTTT	TG	AGTTT	CT	TGG	12987	
Qy	5458	AA	TCTA	AT	GTCT	TAAT	TGG	CAAT	CAT	AAAA	5517	
Db	12988	AA	TCTA	AT	GTCT	TAAT	TGG	CAAT	CAT	AAAA	13047	
Qy	5518	GT	GGTTCT	TG	CTAG	TG	CCGGT	ACT	GT	TG	5577	

13048	GTGGTTTCTGCTAGTCCCGGTACTGTTGTCAGGGGCGCTGTGAGATGCCCCAGTTCCCTGA	13107
5578	AAGAAATGAAAGGCCAGTTTACCGGTAGGTGGTGTGAAACAATCGGCTAGATCATCAGG	5637
13108	AAGAATGAAAGGCCAGTTTACCGGTAGGTGGTGTGAAACAATCGGCTAGATCATCAGG	13167
5638	CAGGACAGAAATGCTCGCTGTGGTGGGAGCACCCAGCTTGGCTTGAAGTTCTCGTTCT	5697
13168	CAGGACAGAAATGCTCGCTGTGGTGGGAGCACCCAGCTTGGCTTGAAGTTCTCGTTCT	13227
5698	ACCACGCGTGTGTTGTGACCAAAATATGAGTTGCTTAAACCTTCTTGTCTACTATTCC	5757
13228	ACCACGCGTGTGTTGTGACCAAAATATGAGTTGCTTAAACCTTCTTGTCTACTATTCC	13287
5758	CTGTTTGC AAAATGTTTCAATTGACCCCTGCTTCCACTCCCAAGGACAAATTTCAACAGC	5817
13288	CTGTTTGC AAAATGTTTCAATTGACCCCTGCTTCCACTCCCAAGGACAAATTTCAACAGC	13347
5818	CTATTGTGTA AAGATCACAGTCTCTTTAAAAATAATAACTGTAAAGTCAGAGGTGATGCT	5877
13348	CTATTGTGTA AAGATCACAGTCTCTTTAAAAATAATAACTGTAAAGTCAGAGGTGATGCT	13407
5878	TGAAAGACAGGAAACCAAGTAGATGGAATGTCATGCTCTTGTGTTCTAAAGAAAAGGC	5937
13408	TGAAAGACAGGAAACCAAGTAGATGGAATGTCATGCTCTTGTGTTCTAAAGAAAAGGC	13467
5938	ATTTTCATAGCTTTTGGATATGACGCAACATACCAATAAATCCTTGACACATAGTTGGAGT	5997
13468	ATTTTCATAGCTTTTGGATATGACGCAACATACCAATAAATCCTTGACACATAGTTGGAGT	13527
5998	CGGAAATGCAACCAAGCCAGTTATAAACCCAGCTAGTTTGGTATGATTGTAAGAAAA	6057
13528	CGGAAATGCAACCAAGCCAGTTATAAACCCAGCTAGTTTGGTATGATTGTAAGAAAA	13587
6058	AAAAGCTGCCAATCTGTAATTTGGGGAATGATTTTCTTAAACTTATATTCTTAGTAG	6117
13588	AAAAGCTGCCAATCTGTAATTTGGGGAATGATTTTCTTAAACTTATATTCTTAGTAG	13647
6118	TCCTAGATTATCATATTGTACTATCATCTCGTGGCTTTTAAAGACTTAAAGAACCAAGTA	6177
13648	TCCTAGATTATCATATTGTACTATCATCTCGTGGCTTTTAAAGACTTAAAGAACCAAGTA	13707
6178	AAATTTTTTTCTTTCTTTAGACATATATAGATCATCAAGGGTGTCTGCTTACAGGTG	6237
13708	AAATTTTTTTCTTTCTTTAGACATATATAGATCATCAAGGGTGTCTGCTTACAGGTG	13767
6238	GATAGTGATATGATCTACAGTGAGGGGACATTTATTTTAAACTTTAAACATTCATGTGTTT	6297
13768	GATAGTGATATGATCTACAGTGAGGGGACATTTATTTTAAACTTTAAACATTCATGTGTTT	13827
6298	TGGGGGTGGTATTTTAAAGGACGACCTCTGATGTCTTTTGGAGGGCTGGTGTGCTTT	6357
13828	TGGGGGTGGTATTTTAAAGGACGACCTCTGATGTCTTTTGGAGGGCTGGTGTGCTTT	13887
6358	GAAAGTCTGTCTCTCTCCAGTGGACTCTAACTTTCTCTGATGCACTGAGACACATTTGT	6417
13888	GAAAGTCTGTCTCTCTCCAGTGGACTCTAACTTTCTCTGATGCACTGAGACACATTTGT	13947
6418	CCTATTGCTCGAGAAACTAAAGCCAAACACGTGTCATCTGGGGACAGGTTTTCAATTTGT	6477
13948	CCTATTGCTCGAGAAACTAAAGCCAAACACGTGTCATCTGGGGACAGGTTTTCAATTTGT	14007
6478	CAGATCTCTTTCCGCCACATGATGTTTCTGGACCAATACAGCTGCTTCCAAAACCTTG	6537
14008	CAGATCTCTTTCCGCCACATGATGTTTCTGGACCAATACAGCTGCTTCCAAAACCTTG	14067
6538	CTAAATTTTGCACAGCTTTCTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTCTGTGAA	6597
14068	CTAAATTTTGCACAGCTTTCTCCTAGGTGCTTGCCCAATGCCAGACTTTCTTTTCTGTGAA	14127
6598	GATTAAGTTGTGCTGTGCTCCCTCTAGTCGGTCAGTTGTTTAAATCTTAAACCTTAAACGGCT	6657
14128	GATTAAGTTGTGCTGTGCTCCCTCTAGTCGGTCAGTTGTTTAAATCTTAAACCTTAAACGGCT	14187

Db 181 GGAAACCCAGGGCTGTAGAGGGGTGAGTGACAAGTCTTTACAAAGTGGCCCTTATTTCCAATC 240
QY 2714 CAGAAATTCGCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 2773
Db 241 CAGAAATTCGCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300
QY 2774 AACTCAATCCCTCTTAATGTACATGGATGGCCAGAGGATGATGGCAGCTCTCTTTGGCAG 2833
Db 301 AACTCAATCCCTCTTAATGTACATGGATGGCCAGAGGATGATGGCAGCTCTCTTTGGCAG 360
QY 2834 TCCGATGGAGATGGAGGATGCCCTTGTCAATGAAGGGACCGCTGTTTCCATTCGGAGC 2893
Db 361 TCCGATGGAGATGGAGGATGCCCTTGTCAATGAAGGGACCGCTGTTTCCATTCGGAGC 420
QY 2894 TACACAGAAAAATGTCAATCGAAGGGGTATATGCCCTTGGATGTGATTTCTG 2953
Db 421 TACACAGAAAAATGTCAATCGAAGGGGTATATGCCCTTGGATGTGATTTCTG 480
QY 2954 CAGCCAGACCTTCACACATTCAGAGACCTTAATAACATGCTTAATGCAACACCGGCC 3013
Db 481 CAGCCAGACCTTCACACATTCAGAGACCTTAATAACATGCTTAATGCAACACCGGCC 540
QY 3014 TACCCTCTGTGAACACGACGATTTCTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 3073
Db 541 TACCCTCTGTGAACACGACGATTTCTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 600
QY 3074 AAGTCAAGTGGCAACAGAACCTCCCAAGGAAAGAAATTCGAAGGAAATGAATTTAGCTG 3133
Db 601 AAGTCAAGTGGCAACAGAACCTCCCAAGGAAAGAAATTCGAAGGAAATGAATTTAGCTG 660
QY 3134 TCAGGTATGTGGCGACACATTTAGAGTGCCTTTTGTATGTGAGATCCACATGAGAACACA 3193
Db 661 TCAGGTATGTGGCGACACATTTAGAGTGCCTTTTGTATGTGAGATCCACATGAGAACACA 720
QY 3194 CAAAGATCTTTTCACTTACGGGTGTAAACATGTGCGGAAGAAATCAAGGACCTTGGTT 3253
Db 721 CAAAGATCTTTTCACTTACGGGTGTAAACATGTGCGGAAGAAATCAAGGACCTTGGTT 780
QY 3254 TCTTAAATATCATGCGGACACATAATGGCAAAATCGGGGGCCAGAGCAAACTGAGCA 3313
Db 781 TCTTAAATATCATGCGGACACATAATGGCAAAATCGGGGGCCAGAGCAAACTGAGCA 840
QY 3314 AGGCTTGGAGATGAGTCCAGCAACCATCAACAGATCGTCCAGGTGACGCGGCGCAGAG 3373
Db 841 AGGCTTGGAGATGAGTCCAGCAACCATCAACAGATCGTCCAGGTGACGCGGCGCAGAG 900
QY 3374 CATCTCCTCTCTTACAAATCTGCATGTTTGTGGCTTCTCTATTTCAAATAAAGAAAG 3433
Db 901 CATCTCCTCTCTTACAAATCTGCATGTTTGTGGCTTCTCTATTTCAAATAAAGAAAG 960
QY 3434 TCTAATTTGAGCACCGCAAGGTGACACCAAAAAATCTGTTTCGGTACAGAGCGCGCA 3493
Db 961 TCTAATTTGAGCACCGCAAGGTGACACCAAAAAATCTGTTTCGGTACAGAGCGCGCA 1020
QY 3494 GACAGACTCTCCACAGGAGGATGCCGTCCTCGAGGAGGACTCTCCAGTGTGTTCAA 3553
Db 1021 GACAGACTCTCCACAGGAGGATGCCGTCCTCGAGGAGGACTCTCCAGTGTGTTCAA 1080
QY 3554 CTTTGAGACCAAAATCTCACCTGAAACCGGGGAAGACCTGTCCAGATGATCCCTCAGCT 3613
Db 1081 CTTTGAGACCAAAATCTCACCTGAAACCGGGGAAGACCTGTCCAGATGATCCCTCAGCT 1140
QY 3614 CGATCGGTTTACACCTTCAGGCTTGGCAGCTGGCTACCAAGGAAAGTTGCCATTG 3673
Db 1141 CGATCGGTTTACACCTTCAGGCTTGGCAGCTGGCTACCAAGGAAAGTTGCCATTG 1200
QY 3674 CCAAGAGTGAAGGAATCGGCGCAAGAGGGAGCACCGACACGATTCGAGTTCCGA 3733
Db 1201 CCAAGAGTGAAGGAATCGGCGCAAGAGGGAGCACCGACACGATTCGAGTTCCGA 1260
QY 3734 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGAGGCTCTCTCGCAAGAGAAAGAGAA 3793
Db 1261 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGAGGCTCTCTCGCAAGAGAAAGAGAA 1320

QY 3794 GTGCAACACCTCCACGGGAGCGCCTCCGTGACGCGGATCCCAAGTTTACCCAGTAG 3853
Db 1321 GTGCAACACCTCCACGGGAGCGCCTCCGTGACGCGGATCCCAAGTTTACCCAGTAG 1380
QY 3854 CAAAGAGAGCCCACTCACTGCTCCGAGTGGGGCAAAAGCTTTCAGAACTTACCAACAGCT 3913
Db 1381 CAAAGAGAGCCCACTCACTGCTCCGAGTGGGGCAAAAGCTTTCAGAACTTACCAACAGCT 1440
QY 3914 GGTCTTGCACTCCAGGGTCCAAAGAGAACCGGAGGCGCGGAGTGCGCCACCAT 3973
Db 1441 GGTCTTGCACTCCAGGGTCCAAAGAGAACCGGAGGCGCGGAGTGCGCCACCAT 1500
QY 3974 GTCTGTGACGGGAGCGCGGAGCGCTGCTCTGACCTCGCGGCCCTCTGGATGA 4033
Db 1501 GTCTGTGACGGGAGCGCGGAGCGCTGCTCTGACCTCGCGGCCCTCTGGATGA 1560
QY 4034 AAATGGAGCCGTGGATCGAGGGAGAGTGGTTCTCAAGACGGATCTGAGGATGGGCTTCC 4093
Db 1561 AAATGGAGCCGTGGATCGAGGGAGAGTGGTTCTCAAGACGGATCTGAGGATGGGCTTCC 1620
QY 4094 CGAAGGAATCCATCTGGGTAA 4114
Db 1621 CGAAGGAATCCATCTGGATAA 1641

RESULT 4

US-08-731-499-10
; Sequence 10, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION:
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOWBEL, David
; APPLICANT: ROMMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,499
; FILING DATE: 16-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/680,395
; FILING DATE: 15-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 23070-068910
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3186 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: -
LOCATION: 1..3186
OTHER INFORMATION: /note= "ZABC1 Open Reading Frame"
US-08-731-499-10

Query Match 15.6%; Score 1559.6; DB 8; Length 3186;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 1565; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY	8286	AGGTGAAAACCATACAAATGTAATTTTGTGTAATGCTGCGCCAGAGACATCTCT	8345
DB	1482	AGGTGAAAACCATACAAATGTAATTTTGTGTAATGCTGCGCCAGAGACATCTCT	1541
QY	8346	GAGGTATCACTTGGAGAGACATCAAGGAAAAACAAACCGATGTTGCTGCTCAAGTCAA	8405
DB	1542	GAGGTATCACTTGGAGAGACATCAAGGAAAAACAAACCGATGTTGCTGCTCAAGTCAA	1601
QY	8406	GAACGATGTTAAAAATCAGGACACTGAAGATGCACTATTAAACCGTGACAGTGGCAAA	8465
DB	1602	GAACGATGTTAAAAATCAGGACACTGAAGATGCACTATTAAACCGTGACAGTGGCAAA	1661
QY	8466	CAAAAAATTTGAAAGATTTTTTGTGTCGTCGCAAGATGTTACAGCAGTCCACTGCAA	8525
DB	1662	CAAAAAATTTGAAAGATTTTTTGTGTCGTCGCAAGATGTTACAGCAGTCCACTGCAA	1721
QY	8526	GCAGCTTAAGGAGATGCTTCTGTTTTCAGAAATGTTCTGGGCGAGCGTCTCTCAAC	8585
DB	1722	GCAGCTTAAGGAGATGCTTCTGTTTTCAGAAATGTTCTGGGCGAGCGTCTCTCAAC	1781
QY	8586	AGCACAAAGATATCAGGATTTCCATAAAAAATGACGTGATGACAGTCTGATAAAGT	8645
DB	1782	AGCACAAAGATATCAGGATTTCCATAAAAAATGACGTGATGACAGTCTGATAAAGT	1841
QY	8646	GATTAACAAACCTTACCTGACCTTACCTGACCTGTTTAAAAAGAGATCAGCAGTGAAC	8705
DB	1842	GATTAACAAACCTTACCTGACCTTACCTGACCTGTTTAAAAAGAGATCAGCAGTGAAC	1901
QY	8706	TCAGGCAAAATACCTCATCTGTAGAACCAAGGCGGATGTTACCTCTCCGATGGCAG	8765
DB	1902	TCAGGCAAAATACCTCATCTGTAGAACCAAGGCGGATGTTACCTCTCCGATGGCAG	1961
QY	8766	TACCAACCAATACCTTGAAGTTAGCCCCCAAGAGAGAACCGAGACCGCAGCTGACTG	8825
DB	1962	TACCAACCAATACCTTGAAGTTAGCCCCCAAGAGAGAACCGAGACCGCAGCTGACTG	2021
QY	8826	CAGATACAGCCCAAGTGTGGATTTGTACGAAAAACCTTTAAATTTATCGTGGGGCTCT	8885
DB	2022	CAGATACAGCCCAAGTGTGGATTTGTACGAAAAACCTTTAAATTTATCGTGGGGCTCT	2081
QY	8886	TCACAATTGCGCGCAATTTCTTTGAGTAAAGTTTGATTCCAAAGTATCACCTGTCCATT	8945
DB	2082	TCACAATTGCGCGCAATTTCTTTGAGTAAAGTTTGATTCCAAAGTATCACCTGTCCATT	2141
QY	8946	TTGTACCTTCAAGACATTTTATCCAGAAGTTTAAATGATGACACAGAGACTGGAGCATAA	9005
DB	2142	TTGTACCTTCAAGACATTTTATCCAGAAGTTTAAATGATGACACAGAGACTGGAGCATAA	2201
QY	9006	ATACATCTCGAGTTTCAAAAACTGTGAAACAGTCTGCTTTAGAAAGTCAGGTAC	9065
DB	2202	ATACATCTCGAGTTTCAAAAACTGTGAAACAGTCTGCTTTAGAAAGTCAGGTAC	2261
QY	9066	CGGATCCCGCAGCGTGTGCGGAAAAAGATGTCCTCCCTCCCTAGTTTCTGTAACCC	9125
DB	2262	CGGATCCCGCAGCGTGTGCGGAAAAAGATGTCCTCCCTCCCTAGTTTCTGTAACCC	2321
QY	9126	CAAGCCCAAGTCTGCTTTTCCCGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCA	9185
DB	2322	CAAGCCCAAGTCTGCTTTTCCCGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCA	2381
QY	9186	GAGCCCTCTGGGCCAGGCAAGGCCCTCTGTACTTCAGGGATAGACTCTAGCACTTTAGC	9245
DB	2382	GAGCCCTCTGGGCCAGGCAAGGCCCTCTGTACTTCAGGGATAGACTCTAGCACTTTAGC	2441

QY	9246	CCCAAGTAACCTGAAGTCCACAGACCAAGAGAAATGTTGGGGTCCAAAGGGCCCGCCAC	9305
DB	2442	CCCAAGTAACCTGAAGTCCACAGACCAAGAGAAATGTTGGGGTCCAAAGGGCCCGCCAC	2501
QY	9306	CAGGCAACAGCAATCTGAGATGTTTCTAAACCAAGTGTTCCTCCGTCACCGGATTAAGAC	9365
DB	2502	CAGGCAACAGCAATCTGAGATGTTTCTAAACCAAGTGTTCCTCCGTCACCGGATTAAGAC	2561
QY	9366	AAAAAGACCGGAGACAAATTTGAAACCTTCCAGTAGTCTCTTCAGCCCAACCTCGG	9425
DB	2562	AAAAAGACCGGAGACAAATTTGAAACCTTCCAGTAGTCTCTTCAGCCCAACCTCGG	2621
QY	9426	CAGCAGTAACATCAATGTTTCCATCGACTACCCCGCAAGAACGACAGCCCGTGGGCACC	9485
DB	2622	CAGCAGTAACATCAATGTTTCCATCGACTACCCCGCAAGAACGACAGCCCGTGGGCACC	2681
QY	9486	TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCAGCAATATCTGCAGCAGAAATTTGGTGA	9545
DB	2682	TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCAGCAATATCTGCAGCAGAAATTTGGTGA	2741
QY	9546	GCCCTTTCCAAAAAGACTGAAGTCCAGCGTGTGTCCTTTCAGGTTGACCGCCCGGGC	9605
DB	2742	GCCCTTTCCAAAAAGACTGAAGTCCAGCGTGTGTCCTTTCAGGTTGACCGCCCGGGC	2801
QY	9606	CAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGTTCCAGAGGCATCACATCACT	9665
DB	2802	CAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGTTCCAGAGGCATCACATCACT	2861
QY	9666	GTTACCGCAGGACTGTGTGTATCCGTCGACGCGTGTCTCCCAAAACCAAGGTTCTGTAG	9725
DB	2862	GTTACCGCAGGACTGTGTGTATCCGTCGACGCGTGTCTCCCAAAACCAAGGTTCTGTAG	2921
QY	9726	CTCCAGCGAGGTCGATTTCTCCAAATGTCGACTGTTTCAGAGCCCTATGTTGGCTCCGG	9785
DB	2922	CTCCAGCGAGGTCGATTTCTCCAAATGTCGACTGTTTCAGAGCCCTATGTTGGCTCCGG	2981
QY	9786	GCCACTTTACACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	9845
DB	2982	GCCACTTTACACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3041
QY	9846	TGATGAGGGGGT 9859	
DB	3042	TGATGAGTGTGCT 3055	

RESULT 5

US-10-087-192-1709
; Sequence 1709, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1709
; LENGTH: 5252
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1709

Query Match 15.6%; Score 1559.6; DB 13; Length 5252;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1565; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 8286 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGAGCCCGCAGAGACATCTCT 8345
Db |||||||
QY 1506 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGAGCCCGCAGAGACATCTCT 1565
Db |||||||
QY 8346 GAGGTATCATCTGGAGAGACATCACAGGAAAAACAAACCGGATGTTGCTGCTGAAGTCAA 8405
Db |||||||
QY 1566 GAGGTATCATCTGGAGAGACATCACAGGAAAAACAAACCGGATGTTGCTGCTGAAGTCAA 1625
Db |||||||
QY 8406 GAACGATGTTAAAAATCAGGACACTGAAGATGCACTATTAAACCGCTGACAGTGCACAAAC 8465
Db |||||||
QY 1626 GAACGATGTTAAAAATCAGGACACTGAAGATGCACTATTAAACCGCTGACAGTGCACAAAC 1685
Db |||||||
QY 8466 CAAAAATTTGAAAAAGATTTTTTGTGATGTCGCCAAAGATGTTACAGGAGTGCACCTGCAAA 8525
Db |||||||
QY 1686 CAAAAATTTGAAAAAGATTTTTTGTGATGTCGCCAAAGATGTTACAGGAGTGCACCTGCAAA 1745
Db |||||||
QY 8526 GCAGCTTAAGGAGATGCCCTTCTGTTTTCAGAAATGTTCTGGGACGGCTGCTCTCACC 8585
Db |||||||
QY 1746 GCAGCTTAAGGAGATGCCCTTCTGTTTTCAGAAATGTTCTGGGACGGCTGCTCTCACC 1805
Db |||||||
QY 8586 AGCACACAAAGATACTCAGGATTTCCATAAAAAATCAGCTGATGACAGTGTGATAAAGT 8645
Db |||||||
QY 1806 AGCACACAAAGATACTCAGGATTTCCATAAAAAATCAGCTGATGACAGTGTGATAAAGT 1865
Db |||||||
QY 8646 GAATAAAAACCTTACCCCTGCTTACCTGGACCTGTTTAAAAAGAGATGACGAGTGTGAAAC 8705
Db |||||||
QY 1866 GAATAAAAACCTTACCCCTGCTTACCTGGACCTGTTTAAAAAGAGATGACGAGTGTGAAAC 1925
Db |||||||
QY 8706 TCAGGCAATTAACCTCATCTGTAGAACCAAGGCGATGTTACTCTCTCCGATGGGCGAG 8765
Db |||||||
QY 1926 TCAGGCAATTAACCTCATCTGTAGAACCAAGGCGATGTTACTCTCTCCGATGGGCGAG 1985
Db |||||||
QY 8766 TACCACCCATAACCTTGAAGTTAGCCCAAGAGAGCAAAACGGAGACCGGAGCTGACTG 8825
Db |||||||
QY 1986 TACCACCCATAACCTTGAAGTTAGCCCAAGAGAGCAAAACGGAGACCGGAGCTGACTG 2045
Db |||||||
QY 8826 CAGATACAGGCCAAGTGTGGATGTGTCAGAAAAACCTTTAAATTTATTCGTTGGGGGCTCT 8885
Db |||||||
QY 2046 CAGATACAGGCCAAGTGTGGATGTGTCAGAAAAACCTTTAAATTTATTCGTTGGGGGCTCT 2105
Db |||||||
QY 8886 TCACAAATCCCGCGCAATTTCTTGAGTAAAGTTTGTTCGAAGTATCACCTGTCCATT 8945
Db |||||||
QY 2106 TCACAAATCCCGCGCAATTTCTTGAGTAAAGTTTGTTCGAAGTATCACCTGTCCATT 2165
Db |||||||
QY 8946 TTGTACTCTCAAGACATTTTATCCAGAGTTTTTAAATGATGCACACGAGACTGGAGCATAA 9005
Db |||||||
QY 2166 TTGTACTCTCAAGACATTTTATCCAGAGTTTTTAAATGATGCACACGAGACTGGAGCATAA 2225
Db |||||||
QY 9006 ATACAAATCTGAGCTTCAFAAAAACCTGTGCAAAACAAAGTCTTGTAGAAAGTGCAGGTAC 9065
Db |||||||
QY 2226 ATACAAATCTGAGCTTCAFAAAAACCTGTGCAAAACAAAGTCTTGTAGAAAGTGCAGGTAC 2285
Db |||||||
QY 9066 CCGATGCCCGCAGCGTTCCTGGGAAAGATGTCCTCCCTCCCTAGTTTCTGTAAACC 9125
Db |||||||
QY 2286 CCGATGCCCGCAGCGTTCCTGGGAAAGATGTCCTCCCTCCCTAGTTTCTGTAAACC 2345
Db |||||||
QY 9126 CAAAGCCCAAGTCTGCTTTCCCGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCA 9185
Db |||||||
QY 2346 CAAAGCCCAAGTCTGCTTTCCCGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCA 2405
Db |||||||
QY 9186 GAGCCCTCTGGCCAGGCAAGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGC 9245
Db |||||||
QY 2406 GAGCCCTCTGGCCAGGCAAGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGC 2465
Db |||||||
QY 9246 CCCAAGTAACTGAAGTCCACAGACACAGCAGATGCGGGTCCAGGGGCGCCAC 9305
Db |||||||
QY 2466 CCCAAGTAACTGAAGTCCACAGACACAGCAGATGCGGGTCCAGGGGCGCCAC 2525
Db |||||||
QY 9306 CAGGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCCCTGCAACCGGATAAGAC 9365
Db |||||||
QY 2526 CAGGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCCCTGCAACCGGATAAGAC 2585
Db |||||||

QY 9366 AAAAGACCCGAGACAAATTTGAAACCTTCTCAGTAGTCTCTTCCAGCCACCCCTCGG 9425
Db |||||||
QY 2586 AAAAGACCCGAGACAAATTTGAAACCTTCTCAGTAGTCTCTTCCAGCCACCCCTCGG 2645
Db |||||||
QY 9426 CAGCAGTAACTCAATGTTCCATCGACTACCCCGCAGGAAACGACAGCCCGTGGGCACC 9485
Db |||||||
QY 2646 CAGCAGTAACTCAATGTTCCATCGACTACCCCGCAGGAAACGACAGCCCGTGGGCACC 2705
Db |||||||
QY 9486 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCAGCAATATCTGAGCAGAGAAATTTGGTGA 9545
Db |||||||
QY 2706 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCAGCAATATCTGAGCAGAGAAATTTGGTGA 2765
Db |||||||
QY 9546 GCGCCCTTCCAAAAGAGACTGAAGTCCAGCGTGGTTGCCCTTGCAGTTGACCGCCGGGC 9605
Db |||||||
QY 2766 GCGCCCTTCCAAAAGAGACTGAAGTCCAGCGTGGTTGCCCTTGCAGTTGACCGCCGGGC 2825
Db |||||||
QY 9606 CAATTACAGAAAGAGGCTATGACCTTCCCAAGTACCATATGGTTCAGAGGCATCACATCACT 9665
Db |||||||
QY 2826 CAATTACAGAAAGAGGCTATGACCTTCCCAAGTACCATATGGTTCAGAGGCATCACATCACT 2885
Db |||||||
QY 9666 GTTACCGCAGGACTGTGTGATCCGTCGAGGCGTGCCTCCCAAAACCAAGGTTCTCTGAG 9725
Db |||||||
QY 2886 GTTACCGCAGGACTGTGTGATCCGTCGAGGCGTGCCTCCCAAAACCAAGGTTCTCTGAG 2945
Db |||||||
QY 9726 CTCCAGCAGGCTCGATTTCTCCAAATGTCTGACTGTTTCCAGAGCCCTATGTTGGCTCCGG 9785
Db |||||||
QY 2946 CTCCAGCAGGCTCGATTTCTCCAAATGTCTGACTGTTTCCAGAGCCCTATGTTGGCTCCGG 3005
Db |||||||
QY 9786 GCCACTTTACACTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9845
Db |||||||
QY 3006 GCCACTTTACACTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3065
Db |||||||
QY 9846 TGCATGAGGGCGT 9859
Db |||||||
QY 3066 TGGTGGATGCTAGT 3079
Db |||||||

RESULT 6

US-08-731-499-3
; Sequence 3, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION:
; APPLICANT: GRAY, Joe W
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOWBEL, David
; APPLICANT: ROMMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR
; TITLE OF INVENTION: USES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,499
; FILING DATE: 16-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/680,395
; FILING DATE: 15-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498

REFERENCE/DOCKET NUMBER: 23070-068910
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1507 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: -
LOCATION: 1..1507
OTHER INFORMATION: /note= "cdna clone cc49 of 6-7kb
transcript with homology to C2H2 zinc
finger genes"
US-08-731-499-3

Query Match 13.9%; Score 1387.2; DB 8; Length 1507;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1470; Conservative 3; Mismatches 32; Indels 6; Gaps 6;

QY 2424 CAGGTTGCTGGGATGACCTTCTGCTCAATTGAAACACTCAATTCATATGGAGACAAAGAGC 2483
DB 2 CAGGTTGCTGGGATGACCTTCTGCTCAATTGAAACACTCAATTCATATGGAGACAAAGAGC 61
QY 2484 ACTAATGCTTTGCTGCTGATTCATATTTGAATCGAGGCATTTGGAAACCTGTATGCCCTGT 2543
DB 62 ACTAATGCTTTGCTGCTGATTCATATTTGAATCGAGGCATTTGGAAACCTGTATGCCCTGT 121
QY 2544 TTGTGGAAGAACCCAGTGACACCATCACTGAGCTTCTTAAAGTTTGAAGAGTTAGAGG 2603
DB 122 TTGTGGAAGAACCCAGTGACACCATCACTGAGCTTCTTAAAGTTTGAAGAGTTAGAGG 181
QY 2604 ACTATACACTTTCTTTTGAATTTTAAATATTTGCTGCTGTTTGGAAACCCAGG 2663
DB 182 ACTATACACTTTCTTTTGAATTTTAAATATTTGCTGCTGTTTGGAAACCCAGG 240
QY 2664 GCTGTTAGAGGGGTGAGTGACAAAGTCTTCAAGTGGCTTTATTCCTCAACTCCAGAAATTC 2723
DB 241 ACTGTTAGA-GGGTGAAGTACAGTCTTAC-AAGTGGCTTTATTCCTCAACTCCAGAAATTC 298
QY 2724 CCAACGGAACCTTTGAGATATATGCAATCGAAAGTGACAGAAACATGCAACTCAATCC 2783
DB 299 CCAACGGAACCTTTGAGATATATGCAATCGAAAGTGACAGAAACATGCAACTCAATCC 358
QY 2784 CTCTTAATGATGATGGCCAGAAAGTGAATTTGGAGCTCTTTGGCAGTCCGATGGAG 2843
DB 359 CTCTTAATGATGATGGAT-GGCCAAGAGTGAATTTGGCAGCTCTTTGGCAGTCCGATGGAG 417
QY 2844 ATGGAGGATGCTTTGTCATGAAGGACCGCTGTTGTTCCATTCGAGCTACACAGAA 2903
DB 418 ATGGA-GATGCTTTGTCATGAAGGACCGCTGTTGTTCCATTCGAGCTACACAGAA 476
QY 2904 AAAAAATGTC-ATCCAAATCGAGGGTATATGCTTGGATTTGATGCTTTCGAGCCAGAC 2962
DB 477 AAAAAATGTCATCCGAATCGAGGGATATGCTTGGATTTGATGCTTTCGAGCCAGAC 536
QY 2963 CTTTACACATTCAGAGACCTTTAATAACATGCTTTAATGCAACACCGGCTACCTCTG 3022
DB 537 CTTTACACATTCAGAGACCTTTAATAACATGCTTTAATGCAACACCGGCTACCTCTG 596
QY 3023 TGAACAGAGCTTTCTGGGTTGAACGAGATATCTCAGTCCGCTTGATATAAGTCAAGT 3082
DB 597 TGAACAGAGCTTTCTGGGTTGAACGAGATATCTCAGTCCGCTTGATATAAGTCAAGT 656
QY 3083 GCGAACAGACCTTCCCAAGGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGTATG 3142
DB 657 GCGAACAGACCTTCCCAAGGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGTATG 716
QY 3143 TGGGCAGACATTTAGAGTCCGCTTTTGAATTTGATGATCCATGAGAACACACAAAGATTC 3202

DB 717 TGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAAACACACAAAGATTC 776
QY 3203 TTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTTCAGAGGACCTTTGGTTCTTAAAAA 3262
DB 777 TTTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTTCAGAGGACCTTTGGTTCTTAAAAA 836
QY 3263 TCACATGCGGACACATAATGSCAAATCGGGGGCCAGAAAGCAAACTGCAGCAAGGCTTTGGA 3322
DB 837 TCACATGCGGACACATAATGSCAAATCGGGGGCCAGAAAGCAAACTGCAGCAAGGCTTTGGA 896
QY 3323 GAGTAGTCCAGCAACGATCAACGAGTCTGTCAGGTGTCACGCGGCCGAGAGCATCTCTCTC 3382
DB 897 GAGTAGTCCAGCAACGATCAACGAGTCTGTCAGGTGTCACGCGGCCGAGAGCATCTCTCTC 956
QY 3383 TCCTTACAAATCTGATGCTTTGCTCTTATTTTCCAAATAAAGAAAGTCTTAATTGA 3442
DB 957 TCCTTGAANAATCTGATGCTTTGCTCTTATTTTCCAAATAAAGAAAGTCTTAATTGA 1016
QY 3443 GCACCCAGAGGTGCACACCAAAAAAACTGCTTTTCGGTACCAGCAGCGCGCAGACAGACTC 3502
DB 1017 GCACCCAGAGGTGCACACCAAAAAAACTGCTTTTCGGTACCAGCAGCGCGCAGACAGACTC 1076
QY 3503 TCACAAAGAGGAAATGCCGCTCTCGAGGGAGGACTTCTCTGCACTTTGTTCAACTTGAGACC 3562
DB 1077 TCACAAAGAGGAAATGCCGCTCTCGAGGGAGGACTTCTCTGCACTTTGTTCAACTTGAGACC 1136
QY 3563 AAAATCTCACTCCCTGAAACCGGGGAAGAGCTGTGAGATGTCATCCCTCAGCTCCATCCGTT 3622
DB 1137 AAAATCTCACTCCCTGAAACCGGGGAAGAGCTGTGAGATGTCATCCCTCAGCTCCATCCGTT 1196
QY 3623 CACCACTTTCCAGGCTTGGCAGCTGTACCAAGGAAAGTTGCCATTTGCCAAGAAAGT 3682
DB 1197 CACCACTTTCCAGGCTTGGCAGCTGTACCAAGGAAAGTTGCCATTTGCCAAGAAAGT 1256
QY 3683 GAAGGAATCGGGCAAGAGAGGAGCAACCGAACAAGCAATTCGAGTTCCGAGAGGAGCT 3742
DB 1257 GAAGGAATTTGGGCAAGAGAGGAGCAACCGAACAAGCAATTCGAGTTCCGAGAGGAGCT 1316
QY 3743 TGGAGAAACAAATAAGGGCAGTTGTGAGGCTCTCGAGGCTCTCGAAGAGAGAGAGTGCAGAA 3802
DB 1317 TGGAGAAACAAATAAGAACCAATTTGTGAGGCTCTCGAAGAGAGAGAGTGCAGAA 1376
QY 3803 CTCCCAAGGAGGAGGCTCTCGGTCGAGCGGATCCCAAGTTACCCAGTAGCAAGAGAA 3862
DB 1377 CTCCCAAGGAGGAGGCTCTCGGTCGAGCGGATCCCAAGTTACCCAGTAGCAAGAGAA 1436
QY 3863 GGCCTACTCTGCTCCGAGTGGGCAAAAGCTTTTCAAGACCTTACCAAGCTGCTCTTGA 3922
DB 1437 GGCCTACTCTGCTCCGAGTGGGCAAAAGCTTTTCAAGACCTTACCAAGCTGCTCTTGA 1496
QY 3923 CTCCAGGCTCC 3933
DB 1497 CTCCAGGCTCC 1507

RESULT 7

US-10-684-422-156/c
; Sequence 156, Application US/10684422
; Publication No. US2004022923A1
; GENERAL INFORMATION:
; APPLICANT: ABURATANI, Hiroyuki
; APPLICANT: YAMAMOTO, Shogo
; TITLE OF INVENTION: Human housekeeping genes and human tissue-specific genes
; FILE REFERENCE: 113991
; CURRENT APPLICATION NUMBER: US/10/684,422
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/418,614
; PRIOR FILING DATE: 2002-10-16
; NUMBER OF SEQ ID NOS: 332
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 156
; LENGTH: 124990
; TYPE: DNA


```

; ORGANISM: Homo sapiens
US-10-684-422-156

Query Match      10.1%; Score 1006.8; DB 20; Length 124990;
Best Local Similarity 95.5%; Pred. No. 6.3e-225;
Matches 1083; Conservative 0; Mismatches 37; Indels 14; Gaps 4;

QY 1 CCATCATATTTCTTATTTTGGGCGGAGAGGG-----AGACTTGTCTGTGGCCCA 54
Db 1128 CCATCATATTTCTTATTTTGGGCAAGAGGGGGAGTCAAAAGTCTCGCTCTGTGGCCCA 1069

QY 55 GCGTGA--CCAGTGTGGTGGCTTGGCTCACTGCAACCTCCACCTCTCGGTTCAAGTG 112
Db 1068 GCGTGAATGCAAGTGGCGGATCTTGGCTCACTGCAACCTCCACCTCTCGGTTCAAGTG 1009

QY 113 ATTCCCAATAGCTGGGATTACAGGTGTGTATTACCATGCCAGCTAAATTTTGTATTTT 172
Db 1008 ATTCCCAAGTAGCTGGGATTACAGGTGTGTATTACCATGCCAGCTAAATTTTGTATTTT 949

QY 173 TAGCAGATAAGGGTTTACCATGTTGGCGAGCTGGTCTCCAACCTCTGGCTCATGTG 232
Db 948 TAGCAGATAAGGGTTTACCATGTTGGTCAAGCTGGTCTCCAACCTCTGACCTCAAGTG 889

QY 233 ATCCACCCACTTCGGCTTCCCAAAGCATTTGGGAGTATAGTGTGAGCCACTATACCGTC 292
Db 888 ATCCACCCACTTCGGCTTCCCAAAGCATTTGGGAGTACAGGTGTGAGCCACTATGCCCCG 829

QY 293 CTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTCTAAAGGATGT 352
Db 828 CTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTGGAAAGGATGT 769

QY 353 CAGTAGAGAAGTGGAGTTCGCCAAATTTACAGTTTCACGTATTTAGTCAAGTTTCTAAAT 412
Db 768 CAGTAGAGAAGTGGAGTTCGCCAAATTTACAGTTTCACATATTTAGTCAAGTTTCTAAAT 709

QY 413 ACAGTAATAATGTGAGAGCTGACATAGGAGCTAACTGCTGTTTTTTTTTTTTTTTTT 472
Db 708 ACAGTAATAATGTGAGAGCTGACGTAGGAGCTAACTTAG-----TTTTTGTTTTTTTTT 654

QY 473 TTTCAAATTTCTCACTGAACCTTTGATTTTCTAAATAGGACATTTAAAAAACAACCAAA 532
Db 653 TTTCAAATTTCTCACTGAACCTTTGATTTTCTAAATAGGACATT-AAAAAACAACCAAA 595

QY 533 AACTCCACTATTTGCGCTATTTGCCACTATTTTGATTTTAAAAAATAAGCGTATTTAGCAT 592
Db 594 AACTCCACTATTTGCGCTATTTGCCACTATTTTGATTTTAAAAAATAAGCGTATTTAGCAT 535

QY 593 CTAAAGTAGGAAGGACCTCAAAATAANTGAGTCTTTGTTCTTGGCCAGGAAAAACAGCGT 652
Db 534 CTGCAAGTAGGAAGGCGCTCAAAATAANTGAGTCTTTGTTCTTGGCCAGGAAAAACAGCGT 475

QY 653 TGTGAGAAATTTGATAAATCTGTTTTCTAGGGTATGTCTGTTATTCAGTTTAAACCTTGC 712
Db 474 TGTGAGCAATTTGTTAACTGTTTTCTAGGGTATGTCTGTTATTCAGTTTAAACCTTGC 415

QY 713 TGGGAGCGCTAGCAATTCAGTAAATACCTGTTGTAATAGCAAAATGAACCTTAAAGCTTATG 772
Db 414 TGGGAGCGCTAGCAATTCAGTAAATACCTGTTGTAATAGCAAAATGAACCTTAAAGCTTATG 355

QY 773 TATAGAAACCTAAGTCACTTTCATTTCTGATTTAGCAGAGTAAATTTGAAATTTCTTTTCAAT 832
Db 354 TATAGAAACCTAAGTCACTTTCATTTCTGATTTAGCAGAGTAAATTTGAAATTTCTTTTCAAT 295

QY 833 GTGTAGCTCTATCCCAGAACCAAGAAATTTGGAACTGTAAAGGCGCACTTATAGTTTA 892
Db 294 GTGTAGATCTATCCCAGAACCAAGAAATTTGGAACTGTAAAGGCGCACTTATAGTTTA 235

QY 893 ACCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTGCGAGTGCAACTTGAAGT 952
Db 234 ACCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTGCGAGTGCAACTTGAAGT 175

QY 953 TGTGACTAGAACTCGGGTCTCTGGAGTGTCTTATTATATACACCAAGCTGGTCAACCAG 1012

```

```

Db 174 TGTGACTAGAACTCGGTCTCTGGAGTGTCTTATTATATACACCAAGCTGGTCAACCAG 115
QY 1013 CCATGTGTTGATCTCCCATTTGTGATAGCAACAAGAAAGACTTCAGGACATCTTTCTCT 1072
Db 114 CCATGTGTTGATCTCCCATTTGTGATAGCAACAAGAAAGACTTCAGGACATCTTTCTCT 55
QY 1073 TTACCCCTAATCCTTGATCTGCGAGTCTTATTTAGAAAAGCTTAAATGTTAAAGATC 1126
Db 54 TTACCCCTAATCCTTGATCTTTCAGTCTTATTTAGAAAAGCTTAAATGTTAAAGATC 1

RESULT 8
US-10-087-192-1705
; Sequence 1705, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1705
; LENGTH: 26345
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-087-192-1705

Query Match      8.2%; Score 820; DB 13; Length 26345;
Best Local Similarity 68.1%; Pred. No. 2.6e-181;
Matches 1391; Conservative 0; Mismatches 550; Indels 101; Gaps 14;

QY 2132 AACTTGCATTTTAGATTTAAAGTCGGTAAAGCACCTGATTTAACTGGGATTTTAACTGGAT 2191
Db 9394 AAATTTGGCATTTTATATTTTCAGTGAGCAGAGCTCTGCCTTAAGTCTCCTGACTGGTTGAA 9453

QY 2192 GAAATTCGTATTTAAATAGTGTACTGACTGGATTAATAAGCCCAATGATTTAATTAACAAGC 2251
Db 9454 ATTCTACTCAGTGTAGTATGATCCACTGGGAAATATGCTAATGATTTAATTAACAAC 9513

QY 2252 ACGTTTAAACAGGATCCCTCATATATATTAGTTAAAGTGAAGCAATTTGAATTAGGTACCTTC 2311
Db 9514 ATGTTTAGAGGCTGCATTTATGGATTAGTTAAAGGTGAGGAGTTGACTCTAGTCCCTTT 9573

QY 2312 TCTGCTGCTGGAAGAAGACCGTATGACTCAACCCACACGAGCTTCTCTTCGCTCTGAGTG 2371
Db 9574 T-----GTGAGAAGATCATGTGACTCATAGGCCCAAGCCCTGTTCTCTGAGCCTTCCTC 9626

QY 2372 TAGCTAACCGTTTCGTGTTTTTTTCTCTAGGGTTTGGAAATCCCTGTCTCCAGGTTGC 2431
Db 9627 ACTGAGAC-----TTGTTTTTCTTTCAGATTTTGGAAATCCCTGTCTTCACATTCG 9681

QY 2432 TGGGATTCGACTTCTTGTCTCAATTTGAAACACATCATTTCAATGGAGACAAAGAGAACTAATG- 2490
Db 9682 CGGGATCAACGCTTTGCTCACTGAGCCACTCA-CCAGTGGAGATGGACAGAGCTGTTCG 9740

QY 2491 CTTTGTGCTGATTCATATTTGAATCGAGGCAATGGGAAACCCCTGTATGCTTGTGTTGTGA 2550
Db 9741 CCTCGTGTGACGCACATCTGACTCACAGCCTGGGGCAGCTTGGATGCCCTTGTGTTGTGCA 9800

QY 2551 AAGAACCGATGACACCATCAGCTGAGCTTCTTAAAGTTTCAAGAGTTTGAAGAGCTATAC 2610
Db 9801 AAGAGCCCA---TCTCCATCAGCTGAGCTCTCCCAAGCTCC-----GGAATATATAC 9845

QY 2611 ACTTCTTTTGAACCTTTTATAATAATATTTTCTCTGTTTGTGAAATCCCAAGCTGTTA 2670

```



```
Db 9846 TCGTCCCTTGACCTT-----AGAGTGCTCTGATTTTGACACCCAGGGCTCTTAG 9896
Qy 2671 GAGGGGTGAGTGAC- AAGTCTTCAAGTGGCTTTATTCCAATCCAGAAATTGCC----- 2724
Db 9897 AGGTTCTGAGTGGTGATATCTACAGGGGGCTTATTCTAACTGCAGAGATCTCCAGCAG 9956
Qy 2725 CAACGGAACCTTTGAGATTATATGCAATGCGAAAGTGACAGGAAACATGCCAACTCAATCCC 2784
Db 9957 ATAGTGATTAACCATCTCTGTGACAGCCCTGAAGTGGTTGGCAGCATGCCGACCCAGTCCC 10016
Qy 2785 TCTTAATGTACATGATGGGCCAGAGTATGTCAGCTCTCTGGCAGTCCGATGAGAGA 2844
Db 10017 TCTCTGTGTACATGGAGCGGGCCGGAAGTCTCTCAGCAGTCTCTAGGCTCCAGATGGAGG 10076
Qy 2845 TGGAGGATGCTTGTCAATGAAGGGACCGCTGTGTTTCCATTCCGAGCTACACAGAAA 2904
Db 10077 TGGATGATGTGTGCCCATAAAGGGCGGGTGGCAGTCCCTTCCGAGCTGCTCAGGAGA 10136
Qy 2905 AAAATGTATCCAAATCGAGGGTATATGCTTGGATTTGATGTTCTGAGCCAGACCT 2964
Db 10137 AGAGCATGGCCGTGGCAGAGGGCCACATGCCCTCGGATTTGCATGTTCTGAGCCAGGTCT 10196
Qy 2965 TCACACATTCAGAGACCTTAATAAATCAATCTTAATGCAACACCGGCTTACCTCTGTG 3024
Db 10197 TCTCTAGGCGGAGGATCTCAGTCAGCACGCTGCTGCTGCGAGCACCGGCCACCTCTGCG 10256
Qy 3025 AACGAGCAGTCTTTCGGGTTGAAGCAGATATCTCAGTCCGCTTGATAAAGTCAAGTGC 3084
Db 10257 AGCCAGTGTCTCGGTGTGGAGCCGAGTACCTTAGTCCCTTGATTAAGTCTTGGAGC 10316
Qy 3085 GAACAGAACCTCCCAAGGAAAGAAATGCAAGGA---AAATGAATTTAGCTGTGAGGTAT 3141
Db 10317 CAACAGAGCCAGCATTTGGAAGAGAGTGGCGAAGACCCCGAGGAGTTGAGCTGTGATGTG 10376
Qy 3142 GTGGCAGACATTTAGAGTGCCTTTTTCATGTTGAGATCCACATGAGAAACACAAAGATT 3201
Db 10377 GTGGCAGACATTTCCCAAGTGGCTTTTTCATGTTGAGAGCCACATGAAGAGCATTAAGGACT 10436
Qy 3202 CTTTCACTACGGGTGAACATGTGCGGAAGAGATTCAAGGAGCCTTGTTCTTAAAA 3261
Db 10437 CTTTCACTATGGTGCAGCATGTGCGGAGGAGATTCAAGAGCCGTGTTCTTGAAGA 10496
Qy 3262 ATCATTGCGGACACATAATGGCAAAATCGGGGGCCAGAAACAACTGCACAAAGGCTTGG 3321
Db 10497 ACCACATGCGGACACACAAATGGCAAGTCTGGCACCGAGGAGCAAGTCTCAGCAAGGCATGG 10556
Qy 3322 AGAGTAGTCAGCAACGATCAACAGGTCTGTCAGGTGCGCGGCGGCGAGGATCTCCT 3381
Db 10557 AG---AGTCCAGTCAACCATCAATGAAGTGGTCCAGCGCGCACGCCCTGGGAGCATCTCCA 10613
Qy 3382 CTCCTTACAAAATCTGCATGTTGTGGCTTCTTATTTCCAAATTAAGAAAGTCTAATTG 3441
Db 10614 CGCCCTACAAGATCTGCATGGTCTGCGGCTTCTCTTCCCAATTAAGAGAGCCTCATTTG 10673
Qy 3442 AGCACCGCAAGGTGCAACCAAAAAAATCTGCTTTCGGTATCCAGCAGCGCGCAGACAGACT 3501
Db 10674 AGCACAGCAAGGTTACGCGCAAGAAACTGTGCCAGTGCAGCAAGGTTGCCCTGATG 10733
Qy 3502 CTCACAGAGAGAAATGCGCTCTCGAGGAGAGACTTCTGCAAGTGTGTTCAACTTGAGAC 3561
Db 10734 ACCACCGAGAGGAACCCACGCTCCCGAGGGAAGATTGCTGCAAGTGTGTTTGAACCTTGAGAC 10793
Qy 3562 CAAATCTCACCTGAAACGGGGAAGAGCCTGTGATGTCATCCCTCAGCTCGATCCGT 3621
Db 10794 CCAGTCAACTGAGGTATGACAGTGAAGCCCATGACCTGCATACCTCAGCTTGACCCGT 10853
Qy 3622 TCACCACTTCCAGGCTTGGCAGTGGCTACCAAGGAAAGTGTGCCATTGCTCC---AAG 3678
Db 10854 TCACCACTTACAGGCATGSCAGTTGGCTACCAAGGAAAGTGGCGGTGTCGCCAGGAG 10913
Qy 3679 AAGTGAAGGATCGGGCAAGAGGGAGCAACGCAACGACGATGTCAGTTCGCGAAGG 3738
Db 10914 AGGTGAAGAGTTCAGGCGCAAGAGGAAGCAACAGCAATGACGACTCATGCTCAGAGAAAG 10973
```

```
Qy 3739 AGCTTGGAG-----AAACAA 3753
Db 10974 AGGNACTAGGGGAAATATGGTTGGGGTAAAGCGGAAGGCTCTGGAAGTCCAAACAA 11033
Qy 3754 ATAAGGCACTTGTGCAAGGCTCTGCAAGAGAAAGAAAGTGAACACACTCCACGCGC 3813
Db 11034 GTAAAAGCAGTTCTCCAGGTCTCTCCCAAGACAAAGAGACCTTAGACATGCTAATAGTG 11093
Qy 3814 AAGGCCCTCCGTGGAGCGGATCCCAAGTTACCACTAGCAAGGAGAGCCCACTCACT 3873
Db 11094 AAGTGCCCTCTGGGATAGTGACCCCAAGTTGTCAGTAGCAAGAGAGCCCACTCACT 11153
Qy 3874 GCTCCAGAGTCGGCAAGCTTTTCAGAACCTTACCAACAGCTGCTTTCGACTCCAGGGTCC 3933
Db 11154 GTTCTGAGTGCAGAAAGCCTTCAGGACATACCAACAGCTGCTCTGCACTCGAGGGTGC 11213
Qy 3934 ACAAGAGAACCCGAGGCGCGGAGTGCGCCACCATGTCTGTGGAACGAGGAGGAGC 3993
Db 11214 ACAGGAAGGACAGGAGGACTGATGCCCTGTGACCCACCATGTGTGATGCAAGGCAGC 11273
Qy 3994 CGGGAGCTGTTCTCTGACCTCGCCGCTCTGATGAAATGGAAGCGCTGATCGAG 4053
Db 11274 CTGGGACCTGCTCCCGACCTCAGCACCATCTGGAAGACAGTGGGGCCGAGGACCGA- 11332
Qy 4054 GCGAAGTGTCTGAAGACGATCTGAGATGGGCTTCCCGAAGGAATCATCTGGGTA 4113
Db 11333 --GAAGGGGCTCTGAAGACGGTCTGAGGATGGAATCTCCCTGACGGGCTCCATTGGGTG 11390
Qy 4114 AG 4115
Db 11391 AG 11392
```

RESULT 9
US-10-087-192-1706
; Sequence 1706, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1706
; LENGTH: 3016
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-087-192-1706

Query Match 7.3%; Score 726.4; DB 13; Length 3016;
Best Local Similarity 70.2%; Pred. No. 8.1e-160;
Matches 1097; Conservative 0; Mismatches 411; Indels 54; Gaps 7;

```
Qy 8286 AGGTGAAAACCATACAAATGTAATTTTGTGAATATGTCAGAGCCGAGAGACATCTCT 8345
Db 1503 AGGTGAAAACCATACAAATGTAATTTTGTGAATATGTCAGAGCCGAGAGACATCTCT 1562
Qy 8346 GAGGTATCACTTGGAGAGACATCAAGAGGAAAAAACAAC---CGATGTTGCTGCTGAAGT 8402
Db 1563 GAGGTACCACTTGGAGAGACATCAAGAGAGAGAGCGGTGATGCTGCTGCTGATGTC 1622
Qy 8403 CAAGAAACGATGTTAAATAATCAGGACACTGAAGATGCACTATTAA---CCGCTCAGAGTGC 8459
Db 1623 CAAAAGTGAAGGCGGAGCGCAGGAGCGGAGGATGCGCTACTAAACGCTGCTGACAGTGC 1682
```

QY 8460 GCAACCAAAATTTTGAAGATTTTTCATGGTCCAAAGATGTTACAGGAGTCCACC 8519
DB 1683 CAGACCAAAATTTTAAAGAGATTTCTTGTATGGTCCAAAGATGTTAAGGGAAAGCCACC 1742
QY 8520 TCAAAAGACGCTTAAGGAGATGCTTCTGTTTTCAGATGTTCTGGGAGCGCTGTCT 8579
DB 1743 TGCCAAAGACGCTTAAGGAGATGCTTCTGTTTCCAGAGTTC-----TCTC 1790
QY 8580 CTCACAGCACAAAGATATCTAGGATTTCCATAAAATGAGCTGATGACAGTGTGA 8639
DB 1791 ACCAGCACACAGATATCTAGGATTTCCATAAAATGAGCTGAT---AGTGTGA 1847
QY 8640 TAAAGTGAATAAAACCTTACCCCTGTTTACCTGACCTGTTTAAATAAGATATCAGAT 8699
DB 1848 GAAAGGAGGAGAGCCCTGCCCCCTACTTATCTGACATGCAAGAAAGA---AAGCAG 1904
QY 8700 TGAATCTCAGGCAATAAATCTCATCTGTAGAAACCAAGGCGGATGTTACTCTCTCCGA 8759
DB 1905 GGAGCTCAGGCAGCAGCCCTGTCTGACACTAGAGGGGGTGGTCTTTAGCAGGGA 1964
QY 8760 TGGCAGTACCCATCACTTGAAGTTAGCCCCAAAGAGAAAGCAAGGAGACCCGAC 8819
DB 1965 AGCTGG-----CCATAGGAGAGATGGATCAGATGC 1997
QY 8820 TGACTGCAGATACAGCCCAAGTGTGGATGTCACGAAACCTTTAAATTTATCCGTGG 8879
DB 1998 TGACTACAGATACAGCCCGGTGCTGACTGCCAGGACAGGCCCTTTGAATCTATCCCTGG 2057
QY 8880 GCTCTTCAATTTGCCCGCAATTTCTTTGAGTAAAGTTTGAATCCCAAGTATCACCTG 8939
DB 2058 GCGCTCACGCTGCTCTGCAATCTCTTTGAGCAAGTGTCTGATCCCAAGATTTGCCTG 2117
QY 8940 TCCATTTGTACTTCAAGACATTTTATCCAGAAGTTTAAATGATGACACAGACTGA 8999
DB 2118 CCCCTTTGTACTTCAAGACCTTTTATCCGAAGTCTTATGATGACACAGACTGA 2177
QY 9000 GCATAATACATCTCGAGTTTATATAAACTGTGAAACAGTCTTCTGTTAGAGTCG 9059
DB 2178 GCACAGGTACACCTTGCAGCCGACCAAGAACCGGACAGCAAGTCTGTCTGAGAACAG 2237
QY 9060 AGTACCGGATGCCCGCAGCGTTGCTGGGAAAGATGCTCCCTCCCTCTAGTTCTG 9119
DB 2238 GGTACCGGTTGCCCTCCGCTTTGCTGGGAAAGATGCTCCCTCTGCTGGCTGCA 2297
QY 9120 TAAACCAAGCCCAAGTCTGCTTTCCCGCGAGTCCAAATCCCTGCCATCTGCGAAGG 9179
DB 2298 CAAGCCCAAGGCAAGACTGCTTCTCACACACTCGAAGTCCCTGCACTCAGAGAAGC 2357
QY 9180 GAAGCAGACCTCTCGGCGCAGGCAAGGCCCTCTGACTTCAGGATAGACTCTAGCAC 9239
DB 2358 TCGCAGGGGGGCTCGGGGCGCAAGCAAGCAACCCAGACTTCAGGACCAAGCAACAGCAC 2417
QY 9240 TTTAGCCCAAGTAACTGAAGTCCCAAGCACAGACAGAGATGTTGGGGGTCCAAAGGGC 9299
DB 2418 TTTAGCCCAAGTAACTGAAGTCCCAAGCACAGAGTCAACCAATGCTGGGGGCAACAG--- 2474
QY 9300 CGCCACAGGCAACAGCAATCTGAGATGTTTCTTAAACAGTGTTCCTCCCTGCCACCGA 9359
DB 2475 TGCCACAGGCGAGCAGCTGAGAGTGTGTTTCCAAAGGTGGGCTCCCTGCTGTATGA 2534
QY 9360 TAAGACAAAAGCCCGACAAATGAACTCTTCAGTAGTCTCTCTCAGCCAC 9419
DB 2535 TAAGGTGAAGAGACCTTGAGCCAAACTGAACTCCCTACAGGCTCCCGCTCTCAGTCCC 2594
QY 9420 CTTCCGACAGTAACTCAATCAATGGTTCCATCGACTACCCCGCAAGAACAGACAGCCG 9479
DB 2595 CTTACAGATTAATAGCAAGGTTCTGTTGAGTATCCGTTGAGGTTGACGGCCCATG 2654
QY 9480 GGCACCTCCGGGAGAGACTATTTCTGTAATCGGAGTCCGACGAATATCTGACAGAAAT 9539
DB 2655 GGCACAGCAAGGAGAGACTACTACTGCCATCGGAATTTCTGGCAGTCCGCGCAGCAGATA 2714

QY 9540 TGTGTAGCCCTTCCAAAAGACTGAAGTCCAGCGTGGTGGCCCTTGACGTTGACGAGCC 9599
DB 2715 CAGTGAGCCATCCCAAAGACTCAAGTCCAGTCCAGTGGTGGTCCCTGGACACAGAGATGC 2774
QY 9600 CGGGCCCAATACAGAAGAGGCTATGACTTTCCCAAGTACCATATGGTTCAGAGGATCAC 9659
DB 2775 AGGACCAATGCGAGAAGGGCTTTGAGTCCCAAGTACCATATGGTTCAGGAGCATCAC 2834
QY 9660 ATCACTGTTACCGCAGGACTGTGTATCCGTCGAGCGGCTGCCCTCCCAACCAAGTT 9719
DB 2835 CTCTTGTACCAACAGAGATGTGTGCGCCCAACCGCTGTGTCGCCCAAAAGCCGTTT 2894
QY 9720 CTTGAGCTCCAGCAGAGTCTGATTTCTCCAAATGTGCTGACTGTTCAAGAGCCCTATGGTGG 9779
DB 2895 CTTGAGCCCTGGGAGGTGGAGTCAACCAAGTGTGTGGCTGTGCAAGAGCCCTACAGTGC 2954
QY 9780 CTCGGGGCACTTTACACTTGTGTGCTGCTGTGTAGTCAGCATCCAGCTCGACGTTAGA 9839
DB 2955 CTCGAGCCCTGTATACCTGTGGACCCGTCGTCGACACGAGGAGGAGCCAGCCCTTGA 3014
QY 9840 AG 9841
DB 3015 AG 3016

RESULT 10
US-10-027-632-104193
; Sequence 104193, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1998-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 104193
; LENGTH: 707
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-104193

Query Match 6.2%; Score 617.2; DB 13; Length 707;
Best Local Similarity 94.4%; Pred. No. 1.8e-134;
Matches 674; Conservative 1; Mismatches 29; Indels 10; Gaps 3;

QY 39 CTTGCTGTGTTGCCAGGCTGGA--CCAGTGTGCGATCTTGCTCAGTGCACCTCCAC 96
DB 2 CTCGCCCTGTGTTGCCAGGCTGGAATGCAATGTCAGTGGCAGTCTTGCTCAGTGCAC 61
QY 97 CTCCTGGGTTCAAGTGATTCCCAAATAGCTGGGATTTACAGGTGTGTATTACCATGCCAG 156
DB 62 TTCTTGGGTTCAAGTGATTTCCCAAGTAGCTGGGATTTACAGGTGTGTATTACCATGCCAG 121
QY 157 CTAATTTTGTATTTTATAGCAGATAGGGGTTTCAACATGTTGGCAGGCTGGTCTCCAA 216
DB 122 CTAATTTTGTATTTTATAGCAGATAGGGGTTTCAACATGTTGGCAGGCTGGTCTCCAA 181

```
Qy 217 CTCCTGCGCTCATGTATCCACCACCTTCGGCTTCCCAAAGCAATGGGAGTAGGTGTG 276
Db 182 CTCCTGCGCTCATGTATCCACCACCTTCGGCTTCCCAAAGCAATGGGAGTAGGTGTG 241
Qy 277 AGCCACTATACCGCTCTCATCATATATTTCTAATCCCGAGATGTAGAGCTGGTGTCTC 336
Db 242 AGCCACTATACCGCTCTCATCATATATTTCTAATCCCGAGATGTAGAGCTGGTGTCTC 301
Qy 337 TTTTCTTAAAGGATGTCAAGTAGAGAGTGGAGTTCCTCCCAAATTTACAGTTTTCAGTATTA 396
Db 302 TTTTGGAAAGGATGTCAAGTAGAGAGTGGAGTTCCTCCCAAATTTACAGTTTTCACATATTA 361
Qy 397 GTCAAGTTTCTAAATAACAGTAATAATGTTGAGAGCTGACATAGGAGTAACCTTGGTTTT 456
Db 362 GTAAAGTTTCTAAATAACAGTAATAATGTTGAGAGCTGACATAGGAGTAACCTTGGTTTT 417
Qy 457 TTTTCTTAAAGGATGTCAAGTAGAGAGTGGAGTTCCTCCCAAATTTACAGTTTTCAGTATTA 516
Db 418 ---TTTCTTAAAGGATGTCAAGTAGAGAGTGGAGTTCCTCCCAAATTTACAGTTTTCACATATTA 474
Qy 517 AAAAAAACCACCAAAAGTCCACTATTTGCTATTTGCTATTTGCTATTTGCTATTTGCTATTTGCTATTT 576
Db 475 -AAAGAAACCAAAAGTCCACTATTTGCTATTTGCTATTTGCTATTTGCTATTTGCTATTTGCTATTT 533
Qy 577 AAGCGTATTTAGCATCTAAAGTAGGAGGACCTCAAAATAAATGAGTCTTTCTTCTTGG 636
Db 534 AAGCGTATTTAGCATCTAAAGTAGGAGGACCTCAAAATAAATGAGTCTTTCTTCTTGG 593
Qy 637 CCAGGAAACACCGGTTGTCAGAAATTTGATACTGTTTTCTAGGGTATGTGCTGTATT 696
Db 594 CCAGGAAACACCGGTTGTCAGAAATTTGATACTGTTTTCTAGGGTATGTGCTGTATT 653
Qy 697 CAGTAAACCTTGGCTGGAGCTAGCATTCAGTAAATCTGTTGGAATAAGC 750
Db 654 CAGTAAACCTTGGCTGGAGCTAGCATTCAGTAAATCTGTTGGAATAAGC 707
```

RESULT 11

```
US-10-027-632-325113
; Sequence 325113, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 325113
; LENGTH: 707
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-325113
```

Query Match 6.2%; Score 617.2; DB 13; Length 707;
Best Local Similarity 94.4%; Pred. No. 1.8e-134;

```
Matches 674; Conservative 1; Mismatches 29; Indels 10; Gaps 3;
Qy 39 CTCCTGCTGTGTCGCCAGGCTGGA--CCAGTGGTGGATCTTGGCTCACTGCACCTCCAC 96
Db 2 CTCGCCCTGTGTGCCAGGCTGGAAATGAGTGGGACGATCTTGGCTCACTGCTACCTCCAC 61
Qy 97 CTCCTGGGTTCAAGTGATTCCTCAAAATAGCTGGGATTTACAGGTGTGTATTACCATGCCAG 156
Db 62 TTCTCGGTTCAAGTGATTTCCCAAGTAGCTGGGATTTACAGGTGTGTATTACCATGCCAG 121
Qy 157 CTAAATTTTGTATTTTAGCAGATAAGGGGTTTCCCATGTTGGCCAGGCTGTCTCAA 216
Db 122 CTAAATTTTGTATTTTAGCAGATAAGGGGTTTCCCATGTTGGCCAGGCTGTCTCAA 181
Qy 217 CTCCTGGCTCATGTGATCCACCACCTTGGCTTCCCAAGCATTTGGGAGTATAGGTGTG 276
Db 182 CTCCTGGCTCATGTGATCCACCACCTTGGCTTCCCAAGCATTTGGGAGTATAGGTGTG 241
Qy 277 AGCCACTATACCGCTCTCTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTC 336
Db 242 AGCCACTATACCGCTCTCTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTC 301
Qy 337 TTTTCTTAAAGGATGTCAAGTAGAGAGTGGAGTTCCTCCCAAATTTACAGTTTTCAGTATTA 396
Db 302 TTTTGGAAAGGATGTCAAGTAGAGAGTGGAGTTCCTCCCAAATTTACAGTTTTCACATATTA 361
Qy 397 GTCAAGTTTCTAAATAACAGTAATAATGTTGAGAGCTGACATAGGAGTAACCTTGGTTTT 456
Db 362 GTAAAGTTTCTAAATAACAGTAATAATGTTGAGAGCTGACATAGGAGTAACCTTGGTTTT 417
Qy 457 TTTTCTTAAAGGATGTCAAGTAGAGAGTGGAGTTCCTCCCAAATTTACAGTTTTCAGTATTA 516
Db 418 ---TTTCTTAAAGGATGTCAAGTAGAGAGTGGAGTTCCTCCCAAATTTACAGTTTTCACATATTA 474
Qy 517 AAAAAAACCACCAAAAGTCCACTATTTGCTATTTGCTATTTGCTATTTGCTATTTGCTATTT 576
Db 475 -AAAGAAACCAAAAGTCCACTATTTGCTATTTGCTATTTGCTATTTGCTATTTGCTATTT 533
Qy 577 AAGCGTATTTAGCATCTAAAGTAGGAGGACCTCAAAATAAATGAGTCTTTCTTCTTGG 636
Db 534 AAGCGTATTTAGCATCTAAAGTAGGAGGACCTCAAAATAAATGAGTCTTTCTTCTTGG 593
Qy 637 CCAGGAAACACCGGTTGTCAGAAATTTGATACTGTTTTCTAGGGTATGTGCTGTATT 696
Db 594 CCAGGAAACACCGGTTGTCAGAAATTTGATACTGTTTTCTAGGGTATGTGCTGTATT 653
Qy 697 CAGTAAACCTTGGCTGGAGCTAGCATTCAGTAAATCTGTTGGAATAAGC 750
Db 654 CAGTAAACCTTGGCTGGAGCTAGCATTCAGTAAATCTGTTGGAATAAGC 707
```

RESULT 12
US-10-027-632-104193
; Sequence 104193, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358

```
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 104193
; LENGTH: 707
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-104193
```

```
Query Match
Best Local Similarity 6.2%; Score 617.2; DB 17; Length 707;
Matches 674; Conservative 1; Mismatches 29; Indels 10; Gaps 3;
```

```
QY 39 CTTGCTCTGTTGCCCAGGCTGGA--CCAGTGGTGGGATCTTGGCTCACTGCAACCTCCAC 96
Db 2 CTCGCCCTGTTGCCCAGGCTGGAATGCAGTGGCAGCATCTTGGCTCACTGCTACCTCCAC 61

QY 97 CTCCTGGGTTCAAGTGATTTCCCAATAGCTGGGATACAGGTGTGTATTTACCATGCCAG 156
Db 62 TTCCCTGGGTTCAAGTGATTTCCCAATAGCTGGGATACAGGTGTGTATTTACCATGCCAG 121

QY 157 CTAATTTTGTATTTTAGCAGATAAGGGGTTTCCACATGTTGGCCAGGCTGGTCTCCAA 216
Db 122 CTAATTTTGTATTTTAGCAGATAAGGGGTTTCCACATGTTGGCCAGGCTGGTCTCCAA 181

QY 217 CTCCTGGCTCATGTGATCCACCCACTTCGGGTTCCAAAGCATTTGGGATATAGGTGTG 276
Db 182 CTCCTGGCTCATGTGATCCACCCACTTCGGGTTCCAAAGCATTTGGGATATAGGTGTG 241

QY 277 AGCCACTATACCGTCTCATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTC 336
Db 242 AGCCACTATACCGTCTCATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTC 301

QY 337 TTTTCTTAAAGGATGTCAGTAGAAGTGGAGTTCGCCAAAATTACAGTTTCCAGTATTA 396
Db 302 TTTTCTTAAAGGATGTCAGTAGAAGTGGAGTTCGCCAAAATTACAGTTTCCAGTATTA 361

QY 397 GTCAAGTTTCTAAAAATACAGTAATAATGTTGAGAGCTGACATAGGAGCTAACTTGGTTT 456
Db 362 GTAAAGTTTCTAAAAATACAGTAATAATGTTGAGAGCTGACATAGGAGCTAACTTGGTTT 417

QY 457 TTTTCTTAAAGGATGTCAGTAGAAGTGGAGTTCGCCAAAATTACAGTTTCCAGTATTA 516
Db 418 ---TTTTTTTCTTAAAGTGTACCTTAACCTTTGATTTTGTCTAAATAGGACATT 474

QY 517 AAAAAAACCACCAAACTCCACTATTGCTTATTCGCTAACTATTGATTTTAAAAAAT 576
Db 475 -AAAGAAAACCACCAAACTCCACTATTGCTTATTCGCTAACTATTGATTTTAAAAAAT 533

QY 577 AAGCGTATTTAGCATCTAAAAAGTAGGAAGGACCTCAAAATAAATAGTCTTTGTTCTGG 636
Db 534 AAGCGTATTTAGCATCTAAAAAGTAGGAAGGACCTCAAAATAAATAGTCTTTGTTCTGG 593

QY 637 CCAGGAAAAACAGCGTTGTGAGATTTGATTAAGTCTTTTCTAGGGTATGCTGTATT 696
Db 594 CCAGGAAAAACAGCGTTGTGAGATTTGATTAAGTCTTTTCTAGGGTATGCTGTATT 653

QY 697 CAGTTAAAAACCTTGCCTGGGACGCTAGCATTCAGTAAATACCTTGTGTAATAAGC 750
Db 654 CAGTTAAAAACCTTGCCTGGGACGCTAGCATTCGTTAAATACCTTGTGTAATAAGC 707
```

RESULT 13

```
US-10-027-632-325113
; Sequence 325113, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
```

```
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 325113
; LENGTH: 707
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-325113
```

```
Query Match
Best Local Similarity 6.2%; Score 617.2; DB 17; Length 707;
Matches 674; Conservative 1; Mismatches 29; Indels 10; Gaps 3;
```

QY	39	CTTGTCTCTGTTGCCCAGGCTGGA--CCAGTGGTGGGATCTTGGCTCACTGCAACCTCCAC	96
Db	2	CTCGCCCTGTTGCCCAGGCTGGAATGCAGTGGCAGCATCTTGGCTCACTGCTACCTCCAC	61
QY	97	CTCCTGGGTTCAAGTGATTTCCCAATAGCTGGGATACAGGTGTGTATTTACCATGCCAG	156
Db	62	TTCTCTGGGTTCAAGTGATTTCCCAATAGCTGGGATACAGGTGTGTATTTACCATGCCAG	121
QY	157	CTAATTTTGTATTTTAGCAGATAAGGGGTTTCCACATGTTGGCCAGGCTGGTCTCCAA	216
Db	122	CTAATTTTGTATTTTAGCAGATAAGGGGTTTCCACATGTTGGCCAGGCTGGTCTCCAA	181
QY	217	CTCCTGGCTCATGTGATCCACCCACTTCGGGTTCCAAAGCATTTGGGATATAGGTGTG	276
Db	182	CTCCTGGCTCATGTGATCCACCCACTTCGGGTTCCAAAGCATTTGGGATATAGGTGTG	241
QY	277	AGCCACTATACCGTCTCATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTC	336
Db	242	AGCCACTATGCCCAGCTCATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTC	301
QY	337	TTTCTTAAAGGATGTCAGTAGAAGTGGAGTTCGCCAAAATTACAGTTTCCAGTATTA	396
Db	302	TTTCTTAAAGGATGTCAGTAGAAGTGGAGTTCGCCAAAATTACAGTTTCCAGTATTA	361
QY	397	GTCAAGTTTCTAAAAATACAGTAATAATGTTGAGAGCTGACATAGGAGCTAACTTGGTTT	456
Db	362	GTAAAGTTTCTAAAAATACAGTAATAATGTTGAGAGCTGACATAGGAGCTAACTTGGTTT	417
QY	457	TTTCTTAAAGGATGTCAGTAGAAGTGGAGTTCGCCAAAATTACAGTTTCCAGTATTA	516
Db	418	---TTTTTTTCTTAAAGTGTACCTTAACCTTTGATTTTGTCTAAATAGGACATT	474
QY	517	AAAAAACCACCAAACTCCACTATTGCTTATTCGCTAACTATTGATTTTAAAAAAT	576
Db	475	-AAAGAAAACCACCAAACTCCACTATTGCTTATTCGCTAACTATTGATTTTAAAAAAT	533
QY	577	AAGCGTATTTAGCATCTAAAAAGTAGGAAGGACCTCAAAATAAATAGTCTTTGTTCTGG	636
Db	534	AAGCGTATTTAGCATCTAAAAAGTAGGAAGGACCTCAAAATAAATAGTCTTTGTTCTGG	593
QY	637	CCAGGAAAAACAGCGTTGTGAGATTTGATTAAGTCTTTTCTAGGGTATGCTGTATT	696
Db	594	CCAGGAAAAACAGCGTTGTGAGATTTGATTAAGTCTTTTCTAGGGTATGCTGTATT	653
QY	697	CAGTTAAAAACCTTGCCTGGGACGCTAGCATTCAGTAAATACCTTGTGTAATAAGC	750
Db	654	CAGTTAAAAACCTTGCCTGGGACGCTAGCATTCGTTAAATACCTTGTGTAATAAGC	707

Db

654 CAGITAAAAACSTTGCCTGGACGCTAGCATTCGGTAAATACTTGTTGAATAAGC 707

RESULT 14

US-10-029-386-9711/c
; Sequence 9711, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 9711
; LENGTH: 530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR20.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.94
; OTHER INFORMATION: SWISSPROT HIT: O75362, EVALUOE 2.00e-67
; OTHER INFORMATION: NT HIT: G411421959, EVALUOE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BF237807.1, EVALUOE 0.00e+00
US-10-029-386-9711

Query Match	5.3%	Score 530;	DB 16;	Length 530;
Best Local Similarity	100.0%;	Prod. No. 5.5e-114;		
Matches 530;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	3712	ACAACGACGATTCGAGTTCGCGAAGAGAGCTTCGAGAAACAATAAGGCGAGTTGTGCAG	3771	
Db	530	ACAACGACGATTCGAGTTCGCGAAGAGAGCTTCGAGAAACAATAAGGCGAGTTGTGCAG	471	
QY	3772	GCCTCTCGCAAGAGAAAGAGAAGTGCACCACTCCACGCGGAAGGCCCTCCGTGTGAAGC	3831	
Db	470	GCCTCTCGCAAGAGAAAGAGAAGTGCACCACTCCACGCGGAAGGCCCTCCGTGTGAAGC	411	
QY	3832	CGGATCCCAAGTTACCCAGTAGCAAGAGAGAGCCCACTCACTCCGAGTCGGGCAAG	3891	
Db	410	CGGATCCCAAGTTACCCAGTAGCAAGAGAGAGCCCACTCACTCCGAGTCGGGCAAG	351	
QY	3892	CTTTGAGAACCTACCAACGAGCTGGTCTTGACATCCAGGCTCCACAGAAGGACCGGAGGG	3951	
Db	350	CTTTGAGAACCTACCAACGAGCTGGTCTTGACATCCAGGCTCCACAGAAGGACCGGAGGG	291	
QY	3952	CGGCGCGGAGTCGCGCCACCATGCTCTGTGACGCGGAGCGACCGGGAGACGTGTTCTCCTG	4011	
Db	290	CGGCGCGGAGTCGCGCCACCATGCTCTGTGACGCGGAGCGACCGGGAGACGTGTTCTCCTG	231	
QY	4012	ACCTCGCCGCCCTCTCGGATGAAATGGAGCCGTGGATCGAGGGGAGGTGGTTCCTGAAG	4071	
Db	230	ACCTCGCCGCCCTCTCGGATGAAATGGAGCCGTGGATCGAGGGGAGGTGGTTCCTGAAG	171	
QY	4072	ACGGATCTGAGGATGGGCTTCCGAGGGAATCCATCTGGGTAAAGTGCCTGTCTCGTC	4131	
Db	170	ACGGATCTGAGGATGGGCTTCCGAGGGAATCCATCTGGGTAAAGTGCCTGTCTCGTC	111	
QY	4132	CGGTGCTGTTCCGCCGTGTGTCGTCTGTCTCCCGCTCTCCCGCTCTCTATTCCCATCTCC	4191	
Db	110	CGGTGCTGTTCCGCCGTGTGTCGTCTGTCTCCCGCTCTCCCGCTCTCTATTCCCATCTCC	51	
QY	4192	AGACAAACGCTGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGG	4241	

Db
50 AGACAA CGCTGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGG 1

RESULT 15

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

	Query Match	3.4%	Score 335.8;	DB 13;	Length 469;
	Best Local Similarity	99.4%;	Pred. No. 3.1e-68;		
	Matches 337;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	3776	CTCGAAGAGAAAGAAAGTGC	CAACAC	TCCACGGCAGAGCGCCCT	CCGTGCACGCGGA 3815
DB	19	CTAGCAAGAGAAAGAAAGTGC	CAACAC	TCCACGGCAGAGCGCCCT	CCGTGCACGCGGA 78
QY	3836	TCCCAAGTTACCCAGTAGCA	AGAGGCCA	CTACTGCTTCGAGTGGCGCA	AAAGCTTTT 3895
DB	79	TCCCAAGTTACCCAGTAGCA	AGAGGCCA	CTACTGCTTCGAGTGGCGCA	AAAGCTTTT 138
QY	3896	CAGAACTTACCAACAGCTGG	TTTGAC	TTCAGGGTCCACAAGAGCA	CCGAGGGCCGG 3955
DB	139	CAGAACTTACCAACAGCTGG	TTTGAC	TTCAGGGTCCACAAGAGCA	CCGAGGGCCGG 198
QY	3956	CGCGAGTGCCTCCACCAAT	GTCTGTGTGA	CGGGAGGACGCGGGGAC	GTGTTCTCTGACCT 4015
DB	199	CGCGAGTGCCTCCACCAAT	GTCTGTGTGA	CGGGAGGACGCGGGGAC	GTGTTCTCTGACCT 258

Qy	4016	CGCCGCCCTCTGGATGAAATGGAGCCGTTGGATCGAGGGGAAGTGGTTCTGAAGACGG	4075
Db	259	CGCCGCCCTCTGGATGAAATGGAGCCGTTGGATCGAGGGGAAGTGGTTCTGAAGACGG	318
Qy	4076	ATCTGAGGATGGGCTTCCCGAAGGAATCCATCTGGGTAA	4114
Db	319	ATCTGAGGATGGGCTTCCCGAAGGAATCCATCTGGGTAA	357

Search completed: August 1, 2005, 03:00:17
Job time : 3882 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 12:46:07 ; Search time 1023 Seconds
(without alignments)
15994.885 Million cell updates/sec

Title: US-08-731-499-9_COPY_1_10000
Perfect score: 10000
Sequence: 1 CCATCATATTTCTTATTTT.....ATTCTTAGATACGGCAGTGG 10000

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfileseq.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10000	100.0	10365	4	US-08-892-695-9
2	9509.2	95.1	20022	4	US-09-949-016-12504
3	9509.2	95.1	20022	4	US-09-949-016-12504
4	1639.4	16.4	5632	3	US-09-560-594-3
5	1639.4	16.4	5632	4	US-09-949-016-862
6	1639.4	16.4	5632	4	US-09-949-016-4262
7	1559.6	15.6	3186	4	US-08-892-695-10
8	1387.2	13.9	1507	2	US-08-880-395-3
9	1387.2	13.9	1507	4	US-08-892-695-3
10	600.6	6.0	601	4	US-09-949-016-34699
11	600.6	6.0	601	4	US-09-949-016-34700
12	600.6	6.0	601	4	US-09-949-016-34701
13	600.6	6.0	601	4	US-09-949-016-34702
14	600.6	6.0	601	4	US-09-949-016-151879
15	600.6	6.0	601	4	US-09-949-016-151880
16	600.6	6.0	601	4	US-09-949-016-151881
17	600.6	6.0	601	4	US-09-949-016-151882
18	230.2	2.3	99580	4	US-09-949-016-17411
19	228.8	2.3	119981	4	US-09-949-016-11844
20	228.8	2.3	119982	4	US-09-949-016-13606
21	228.6	2.3	46841	4	US-09-949-016-13466
22	228.4	2.3	601	4	US-09-949-016-63770
23	227.6	2.3	92334	4	US-09-949-016-13920
24	227.6	2.3	92363	4	US-09-949-016-12146
25	227.2	2.3	601	4	US-09-949-016-26018
26	227.2	2.3	601	4	US-09-949-016-74620
27	227.2	2.3	88906	4	US-09-949-016-17468

c	28	226.2	2.3	27727	4	US-09-949-016-15737	Sequence 15737, A
	29	226.2	2.3	101951	4	US-09-949-016-15648	Sequence 15648, A
	30	225.2	2.3	53737	4	US-09-949-016-16197	Sequence 16197, A
	31	225.2	2.3	462589	4	US-09-949-016-12900	Sequence 12900, A
	32	225.2	2.3	476044	4	US-09-949-016-12412	Sequence 12412, A
	33	224.8	2.2	222691	4	US-09-949-016-11762	Sequence 11762, A
	34	224.8	2.2	222697	4	US-09-949-016-15842	Sequence 15842, A
	35	224.2	2.2	298	4	US-09-513-999C-11828	Sequence 11828, A
	36	224.2	2.2	66247	4	US-09-949-016-16009	Sequence 16009, A
c	37	224.2	2.2	390416	4	US-09-949-016-16923	Sequence 16923, A
	38	224	2.2	66804	4	US-09-740-041-3	Sequence 3, Appli
	39	223.6	2.2	161124	4	US-09-949-016-11760	Sequence 11760, A
c	40	223.6	2.2	176006	4	US-09-949-016-16804	Sequence 16804, A
	41	223.6	2.2	253375	4	US-09-949-016-12849	Sequence 12849, A
	42	223.2	2.2	601	4	US-09-949-016-118216	Sequence 118216, A
	43	222.8	2.2	601	4	US-09-949-016-130896	Sequence 130896, A
	44	222.8	2.2	36203	4	US-09-949-016-15430	Sequence 15430, A
c	45	222.8	2.2	50472	4	US-09-949-016-14038	Sequence 14038, A

ALIGNMENTS

RESULT 1
US-08-892-695-9
; Sequence 9, Application US/08892695A
; Patent No. 6808878
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W
; APPLICANT: Collins, Collin
; APPLICANT: Hwang, Soo In
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: GENES FROM THE 20Q13 AMPLICON AND THEIR USES
; FILE REFERENCE: 2500.124US3
; CURRENT APPLICATION NUMBER: US/08/892,695A
; EARLIER FILING DATE: 1997-07-15
; EARLIER APPLICATION NUMBER: 08/785,532
; EARLIER FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/731,499
; EARLIER FILING DATE: 1996-10-16
; EARLIER APPLICATION NUMBER: 08/680,395
; EARLIER FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 10365
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Genomic
; OTHER INFORMATION: Sequence encoding ZABCI
; Patent No. 6808878
US-08-892-695-9

Query Match	100.0%	Score	10000;	DB	4;	Length	10365;
Best Local Similarity	100.0%	Pred.	No. 0;				
Matches	10000;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	CCATCATATTTCTTATTTT	TTTTTGGCGGAGAGGAGACTT	GGCTCTGTGTGCCAGGCTGG	60		
Db	1	CCATCATATTTCTTATTTT	TTTTTGGCGGAGAGGAGACTT	GGCTCTGTGTGCCAGGCTGG	60		
Qy	61	ACCACTGGTGCATCTTGCT	CATGCAACCTCCTCGGTTCA	AGTATTCCTCA	120		
Db	61	ACCACTGGTGCATCTTGCT	CATGCAACCTCCTCGGTTCA	AGTATTCCTCA	120		
Qy	121	ATAGCTGGGATTACAGGTG	TATTACCATGCCAGCTAAT	TTTGTATTTTAGCAGAT	180		
Db	121	ATAGCTGGGATTACAGGTG	TATTACCATGCCAGCTAAT	TTTGTATTTTAGCAGAT	180		
Qy	181	AAGGGGTTTACCATTGTTG	GCGGCTGCTCCAACCTCCT	GGCCTCATGTGATCCACCC	240		

|||||
181 AAGGGTTTCAAGTTGGCCAGGCTGGTCTCCAACTCTCGCCCTCATGTGATCCACCC 240
QY
241 ACTTCGGCTTCCCAAGCAATGGGAGTATAGGTGTGAGCCACTATACCCGTCCTCACATC 300
Db
241 ACTTCGGCTTCCCAAGCAATGGGAGTATAGGTGTGAGCCACTATACCCGTCCTCACATC 300
QY
301 ATATTCTAATCCGAGACTGTAGAGCTGGTGTCTCTTTTCTAAAGGATGTCAGTAGAG 360
Db
301 ATATTCTAATCCGAGACTGTAGAGCTGGTGTCTCTTTTCTAAAGGATGTCAGTAGAG 360
QY
361 AAGTGGAGTTCGCCAAATTTACAGTTTACAGTATTAGTCAAGTCTTCTAAATACAGTAAT 420
Db
361 AAGTGGAGTTCGCCAAATTTACAGTTTACAGTATTAGTCAAGTCTTCTAAATACAGTAAT 420
QY
421 AATGTTGAGAGCTGACATAGGACTAACTTGGTTTTTTTTTTTTTTTTTTTTTTTTTTTCAAT 480
Db
421 AATGTTGAGAGCTGACATAGGACTAACTTGGTTTTTTTTTTTTTTTTTTTTTTTTTTTCAAT 480
QY
481 TCTCACTGAACCTTGTGATTTGCTAAATAAGGACATTAATAAAAAAACAACAAAACTCCAC 540
Db
481 TCTCACTGAACCTTGTGATTTGCTAAATAAGGACATTAATAAAAAAACAACAAAACTCCAC 540
QY
541 TATTGCTATTGCGCACTATTGATTTTAAAAAATAAGCGTATTTTAGCATCTAAAGT 600
Db
541 TATTGCTATTGCGCACTATTGATTTTAAAAAATAAGCGTATTTTAGCATCTAAAGT 600
QY
601 AGGAAGGACCTCAATAAATAGTCTTTGTTCTGGCCAGGGAACACAGCGTTGTGAGAA 660
Db
601 AGGAAGGACCTCAATAAATAGTCTTTGTTCTGGCCAGGGAACACAGCGTTGTGAGAA 660
QY
661 TTTGATAACTGTTTTCTAGGATATGCTGTGTTATTTAGTTAAACCTTGCCTGGGACGC 720
Db
661 TTTGATAACTGTTTTCTAGGATATGCTGTGTTATTTAGTTAAACCTTGCCTGGGACGC 720
QY
721 TAGCATTTCAGTAAATCTTTGTAATAAGCAATGAACCTTAAGCTTCTATGTATAGAAA 780
Db
721 TAGCATTTCAGTAAATCTTTGTAATAAGCAATGAACCTTAAGCTTCTATGTATAGAAA 780
QY
781 CCTAAGTCACTTCACATTCGATTAGCAGAGTAATGGAATATCTTTTCAATGTGTAGCT 840
Db
781 CCTAAGTCACTTCACATTCGATTAGCAGAGTAATGGAATATCTTTTCAATGTGTAGCT 840
QY
841 CTATCCCAAGAACACAGAAATATGGAACTGTAAAGGCCATCTATAGTTTAAACCACTG 900
Db
841 CTATCCCAAGAACACAGAAATATGGAACTGTAAAGGCCATCTATAGTTTAAACCACTG 900
QY
901 CGTTAAATAGATAATAGAAAGATGTGTATGTGGCAGTCACAACTTGAAGGTTGTGACTA 960
Db
901 CGTTAAATAGATAATAGAAAGATGTGTATGTGGCAGTCACAACTTGAAGGTTGTGACTA 960
QY
961 GAATCGGGTCTCTGGAGTGTCTATTATATATACAAAGCTGGTCAACAGCCCATGTGT 1020
Db
961 GAATCGGGTCTCTGGAGTGTCTATTATATATACAAAGCTGGTCAACAGCCCATGTGT 1020
QY
1021 TGATCCTCATTGTGTATAGCAACAAAGAAAGACTTCAGGACATCTTTCTTACCCTTA 1080
Db
1021 TGATCCTCATTGTGTATAGCAACAAAGAAAGACTTCAGGACATCTTTCTTACCCTTA 1080
QY
1081 ATTCCTTGATCTGCAGTCTTATTATAGAAAAGCTTAATGTTTAAAGATCTAGTTTATCAAAA 1140
Db
1081 ATTCCTTGATCTGCAGTCTTATTATAGAAAAGCTTAATGTTTAAAGATCTAGTTTATCAAAA 1140
QY
1141 CTAAGATAACAGGAGTATGAAATTTCTATTCGGAGTGAAGGAGGAGATGTTCC 1200
Db
1141 CTAAGATAACAGGAGTATGAAATTTCTATTCGGAGTGAAGGAGGAGATGTTCC 1200
QY
1201 TTGGCTTCTCTGAGCCTGAGGCTTCTTGTCTTTTAGGAGTAGAGAGGAGGAA 1260
Db
1201 TTGGCTTCTCTGAGCCTGAGGCTTCTTGTCTTTTAGGAGTAGAGAGGAGGAA 1260
QY
1261 AGTAAAGTATGCTTTTGTGTTTTTAAAGTTTACTTTGCTGGGAGTAGTTTGCATGCTTTTG 1320

Db
1261 AGTAAAGTATGCTTTTGTGTTTTTAAAGTTTACTTTGCTGGAGTAGTTTGCATGCTTTTG 1320
QY
1321 GTTTTCTTGGGTGGAAATTAACGTACTTAAGTTTTTAAGTAGTTGGGACTATTTAAAAACAA 1380
Db
1321 GTTTTCTTGGGTGGAAATTAACGTACTTAAGTTTTTAAGTAGTTGGGACTATTTAAAAACAA 1380
QY
1381 TGCCTATCCAAATGTTTGGCATAAAAGCAGAGGGTATTCGCTTTAGAAGTTAATCTTCTC 1440
Db
1381 TGCCTATCCAAATGTTTGGCATAAAAGCAGAGGGTATTCGCTTTAGAAGTTAATCTTCTC 1440
QY
1441 CAGGAGTGAATAATAGCTTTCTAAACCAGAACAGCAGAGCTAAATAAAGTAAATTTCCAC 1500
Db
1441 CAGGAGTGAATAATAGCTTTCTAAACCAGAACAGCAGAGCTAAATAAAGTAAATTTCCAC 1500
QY
1501 CTGGCCAGTGCATGATGTGAAGGTAGATTTAAAAAATGAGAGGCCCATTTTCTGATGA 1560
Db
1501 CTGGCCAGTGCATGATGTGAAGGTAGATTTAAAAAATGAGAGGCCCATTTTCTGATGA 1560
QY
1561 AAGACTAAGCCATGTTGAACAGCCCTGTTGAGGATTTTATTTAAATCTATACATTCAC 1620
Db
1561 AAGACTAAGCCATGTTGAACAGCCCTGTTGAGGATTTTATTTAAATCTATACATTCAC 1620
QY
1621 AAAGAGCTTTGTGTATGTCTTTCCCTATTTGCTTTTGGACTAGGAAGCCCAACCACT 1680
Db
1621 AAAGAGCTTTGTGTATGTCTTTCCCTATTTGCTTTTGGACTAGGAAGCCCAACCACT 1680
QY
1681 GCTTGTGAAGCCAGAAAAGTCGTTGAAGCAAGCTGGGATTTTGAACAGTGGATTTGAGTT 1740
Db
1681 GCTTGTGAAGCCAGAAAAGTCGTTGAAGCAAGCTGGGATTTTGAACAGTGGATTTGAGTT 1740
QY
1741 TCGAATATCCAGTGAACCAAAATATATCAGGGTTCCTTGGCCAAAGATGAGTGACCATTC 1800
Db
1741 TCGAATATCCAGTGAACCAAAATATATCAGGGTTCCTTGGCCAAAGATGAGTGACCATTC 1800
QY
1801 TGAGGTGTTAAGTATTTCTTGAATGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 1860
Db
1801 TGAGGTGTTAAGTATTTCTTGAATGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 1860
QY
1861 ATTTTGTGAACTCTGTATGTGCAAAATCTCTAAGGGGGTGTGGGCACCTTAGATTTCT 1920
Db
1861 ATTTTGTGAACTCTGTATGTGCAAAATCTCTAAGGGGGTGTGGGCACCTTAGATTTCT 1920
QY
1921 TGGATGACAGATTTGTTGTATATGAACAAATTTTAAATGTTTGTATACACTGATTT 1980
Db
1921 TGGATGACAGATTTGTTGTATATGAACAAATTTTAAATGTTTGTATACACTGATTT 1980
QY
1981 AAAATAGTTTACTAAAGTGTTTTAAATTTTTCATCTTAAATTTTACAGTTCTTATAGTCT 2040
Db
1981 AAAATAGTTTACTAAAGTGTTTTAAATTTTTCATCTTAAATTTTACAGTTCTTATAGTCT 2040
QY
2041 TTAGATTTAGGAGGCTGTGTGATGATCCCATCATGTGCATTTTAGTGGCATTTAAATGT 2100
Db
2041 TTAGATTTAGGAGGCTGTGTGATGATCCCATCATGTGCATTTTAGTGGCATTTAAATGT 2100
QY
2101 ATTCAGCTGAATTTAAACAATTTCTGACCTTAAACATTTGACATTTTAGATTCGGTAA 2160
Db
2101 ATTCAGCTGAATTTAAACAATTTCTGACCTTAAACATTTGACATTTTAGATTCGGTAA 2160
QY
2161 AGCACTGATTTAAACTGGATTTTAACTGGATGAAATTTCTGATTTTAAATAGTACTGACT 2220
Db
2161 AGCACTGATTTAAACTGGATTTTAACTGGATGAAATTTCTGATTTTAAATAGTACTGACT 2220
QY
2221 GGATAAATGCCAATGATTTAAACAAGCAGTTTAAACAGGATGCCCTATATATAGT 2280
Db
2221 GGATAAATGCCAATGATTTAAACAAGCAGTTTAAACAGGATGCCCTATATATAGT 2280
QY
2281 TAAAGTGAAGCAATTTGAATAGGTACCTTCTCTGCTGGTGGAAAAGCCGTATGACTC 2340
Db
2281 TAAAGTGAAGCAATTTGAATAGGTACCTTCTCTGCTGGTGGAAAAGCCGTATGACTC 2340
QY
2341 ACCCACACGAGCTTCTCTCTGCTGAGTGTAGTAAACCGTTTCTGTTTTTTTTTCTCT 2400
Db
2341 ACCCACACGAGCTTCTCTCTGCTGAGTGTAGTAAACCGTTTCTGTTTTTTTTTCTCT 2400

QY 2401 AGGTTTGGAAATCCCTGTCTCCAGGTGTCTGGGATTTGACTTCTTCTCAATTGAACA 2460
DB |||||
2401 AGGTTTGGAAATCCCTGTCTCCAGGTGTCTGGGATTTGACTTCTTCTCAATTGAACA 2460
QY 2461 CTCAATTCAGAGACAAAGAGACTAATGCTTTGTCTGATTCATATTTGAATCGAGGC 2520
DB |||||
2461 CTCAATTCAGAGACAAAGAGACTAATGCTTTGTCTGATTCATATTTGAATCGAGGC 2520
QY 2521 ATTGGGAACCTGTATGCCCTTTGTGGAAAGAACCAAGTGACACCACTACTGAGCTTC 2580
DB |||||
2521 ATTGGGAACCTGTATGCCCTTTGTGGAAAGAACCAAGTGACACCACTACTGAGCTTC 2580
QY 2581 TAAAGTTTCSAAGAGTTAGAGACTATACATTTCTTTTCACTTTTATTAATAAATATT 2640
DB |||||
2581 TAAAGTTTCSAAGAGTTAGAGACTATACATTTCTTTTGAACCTTTTATAATAAATATT 2640
QY 2641 TGCTCTGTTTGGAAACCCAGGCTGTAGAGGGGTGAGTGACAAGTCTTACAAGTGC 2700
DB |||||
2641 TGCTCTGTTTGGAAACCCAGGCTGTAGAGGGGTGAGTGACAAGTCTTACAAGTGC 2700
QY 2701 CTTATTCCAACTCCAGAAATGCCCACAGAACTTTTGAGATTATATGCAATCGAAAGTGA 2760
DB |||||
2701 CTTATTCCAACTCCAGAAATGCCCACAGAACTTTTGAGATTATATGCAATCGAAAGTGA 2760
QY 2761 CAGGAAACATGSCAACTCAATCCCTCTTAATGTACATGATGGGCCAGAAAGTGA 2820
DB |||||
2761 CAGGAAACATGSCAACTCAATCCCTCTTAATGTACATGATGGGCCAGAAAGTGA 2820
QY 2821 GCTCTCTTGGCAGTCCGATGGAGATGGAGATGCTTGTCAATGAAGGACCGCTGTTG 2880
DB |||||
2821 GCTCTCTTGGCAGTCCGATGGAGATGGAGATGCTTGTCAATGAAGGACCGCTGTTG 2880
QY 2881 TTCCATTCCGAGCTACACAAGAAATAATGTCAATCGAGGGGTATATGCCCTTGG 2940
DB |||||
2881 TTCCATTCCGAGCTACACAAGAAATAATGTCAATCGAGGGGTATATGCCCTTGG 2940
QY 2941 ATTGCATGTTCTGCAGCAGACCTTCAACATTCAGAAGACCTTAATAACATGTCTTAA 3000
DB |||||
2941 ATTGCATGTTCTGCAGCAGACCTTCAACATTCAGAAGACCTTAATAACATGTCTTAA 3000
QY 3001 TGCAACACCGGCTACCTCTGTGAACAGCAGTCTTTCGGGTGAAGCAGATATCTCA 3060
DB |||||
3001 TGCAACACCGGCTACCTCTGTGAACAGCAGTCTTTCGGGTGAAGCAGATATCTCA 3060
QY 3061 GTCCGCTTGATAAAGTCAAGTGCAGACAACTCCCAAGGAAAGAAATTCAGAGAA 3120
DB |||||
3061 GTCCGCTTGATAAAGTCAAGTGCAGACAACTCCCAAGGAAAGAAATTCAGAGAA 3120
QY 3121 ATGAATTTAGCTGTGAGGTATGTGGCAGACATTTAGAGTGCCTTTTGTGATGTTGAGATCC 3180
DB |||||
3121 ATGAATTTAGCTGTGAGGTATGTGGCAGACATTTAGAGTGCCTTTTGTGATGTTGAGATCC 3180
QY 3181 ACATGAGACACACAAGATTTTTCATTAACCGGTGTAAATGTCGGAAGAGATTC 3240
DB |||||
3181 ACATGAGACACACAAGATTTTTCATTAACCGGTGTAAATGTCGGAAGAGATTC 3240
QY 3241 AGGAGCTTGTGTTCTTAATAATCAATGCGGACATATATGSCAAATCGGGGCCAGAA 3300
DB |||||
3241 AGGAGCTTGTGTTCTTAATAATCAATGCGGACATATATGSCAAATCGGGGCCAGAA 3300
QY 3301 GCAAACTGACGACGAGCTTGGAGGTAGTCCAGCAACGATCAACAGAGGTGCTCCAGGTGC 3360
DB |||||
3301 GCAAACTGACGACGAGCTTGGAGGTAGTCCAGCAACGATCAACAGAGGTGCTCCAGGTGC 3360
QY 3361 ACGCGCGGAGAGCATCTCTCTCTTACAAATCTGCAATGTTGTGGCTTCTTATTTTC 3420
DB |||||
3361 ACGCGCGGAGAGCATCTCTCTCTTACAAATCTGCAATGTTGTGGCTTCTTATTTTC 3420
QY 3421 CAAATAAAGAGTCTAAATTGACACGCAAGGTGACACCAAAATCTGCTTTCGGTA 3480
DB |||||
3421 CAAATAAAGAGTCTAAATTGACACGCAAGGTGACACCAAAATCTGCTTTCGGTA 3480

QY 3481 CCACGACGCGCAGACAGACTCTCCACAGGAGGAATCCGCTCTCTCAGGAGGACTTCC 3540
DB |||||
3481 CCACGACGCGCAGACAGACTCTCCACAGGAGGAATCCGCTCTCTCAGGAGGACTTCC 3540
QY 3541 TGCAGTTGTTCAACTTTGAGACCAAAATCTCACCTGAAAACGGGGAAGAGCTTGTCAAGT 3600
DB |||||
3541 TGCAGTTGTTCAACTTTGAGACCAAAATCTCACCTGAAAACGGGGAAGAGCTTGTCAAGT 3600
QY 3601 GCATCCCTCAGCTCGATCCGTTCCACACCTTTCAGGCTTGGCAGCTGGCTACCAAGGAA 3660
DB |||||
3601 GCATCCCTCAGCTCGATCCGTTCCACACCTTTCAGGCTTGGCAGCTGGCTACCAAGGAA 3660
QY 3661 AAGTTCCCAATTTGCCCAAGAGTGAAGGAATCGGGGCAAGAGGAGCACCGACAACGACG 3720
DB |||||
3661 AAGTTCCCAATTTGCCCAAGAGTGAAGGAATCGGGGCAAGAGGAGCACCGACAACGACG 3720
QY 3721 ATTTCGAGTTCCGAGAAAGGAGCTTTGGAGAAACAATAAGGCGAGTTGTGAGGCTCTCGC 3780
DB |||||
3721 ATTTCGAGTTCCGAGAAAGGAGCTTTGGAGAAACAATAAGGCGAGTTGTGAGGCTCTCGC 3780
QY 3781 AAGAGAAAGAGAGTGCACAACACTCCACGCGCAAGCGCCCTCCGTGGACGCGGATCCCA 3840
DB |||||
3781 AAGAGAAAGAGAGTGCACAACACTCCACGCGCAAGCGCCCTCCGTGGACGCGGATCCCA 3840
QY 3841 AGTTACCCAGTAGCAAGGAGAACCCCACTCACTGCTCCGAGTCCGGCAAAAGCTTTTCAGAA 3900
DB |||||
3841 AGTTACCCAGTAGCAAGGAGAACCCCACTCACTGCTCCGAGTCCGGCAAAAGCTTTTCAGAA 3900
QY 3901 CTTACCAACAGCTGTCTTTCAGCTCCAGGGTCCACAAGAGACCCGAGGCGCGGCGG 3960
DB |||||
3901 CTTACCAACAGCTGTCTTTCAGCTCCAGGGTCCACAAGAGACCCGAGGCGCGGCGG 3960
QY 3961 AGTCGCGCCACCACTGTCTGTGGAGGAGGAGGAGCGCGGAGCTTCTCTGAGCTCGCG 4020
DB |||||
3961 AGTCGCGCCACCACTGTCTGTGGAGGAGGAGGAGCGCGGAGCTTCTCTGAGCTCGCG 4020
QY 4021 CCCCTCTGGATGAAATGGAGCGCTGGATCGAGGGGAAGGTGTTCTGAAGACGGATCTG 4080
DB |||||
4021 CCCCTCTGGATGAAATGGAGCGCTGGATCGAGGGGAAGGTGTTCTGAAGACGGATCTG 4080
QY 4081 AGGATGGGCTTCCGAGGAATTCATCTGGGTAAAGTCCGCTGTCTCTCGGCTCGGCTGT 4140
DB |||||
4081 AGGATGGGCTTCCGAGGAATTCATCTGGGTAAAGTCCGCTGTCTCTCGGCTCGGCTGT 4140
QY 4141 TCGGCTGTGTCTGTCTGTCTCCCGTCTCCCGTCTCTATTTCCCACTCTCCACACACG 4200
DB |||||
4141 TCGGCTGTGTCTGTCTGTCTCCCGTCTCCCGTCTCTATTTCCCACTCTCCACACACG 4200
QY 4201 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTGGTATCACTCT 4260
DB |||||
4201 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTGGTATCACTCT 4260
QY 4261 GTGTAAGTCAATTAACCTCTCAGGGCTTAAATTTCTCATTTCTGTAATAACAGGCTTGA 4320
DB |||||
4261 GTGTAAGTCAATTAACCTCTCAGGGCTTAAATTTCTCATTTCTGTAATAACAGGCTTGA 4320
QY 4321 GTTAAGAGTCTCTGTTTCTGAAATATATATATATATATATATATATATATATATATAT 4380
DB |||||
4321 GTTAAGAGTCTCTGTTTCTGAAATATATATATATATATATATATATATATATATATAT 4380
QY 4381 TCACAAAACACACTTTTAAAAAATAAATTTGTGATCCAGGCCAAAATGCACTGCTTCT 4440
DB |||||
4381 TCACAAAACACACTTTTAAAAAATAAATTTGTGATCCAGGCCAAAATGCACTGCTTCT 4440
QY 4441 TAACTGGGCGGATTTGTTTCCCAATCAGTATCTGGCAATGTCTGGAGGCAATTTGGTGT 4500
DB |||||
4441 TAACTGGGCGGATTTGTTTCCCAATCAGTATCTGGCAATGTCTGGAGGCAATTTGGTGT 4500
QY 4501 CATACTGTGTGTGGGTGTGCTGCTGCACTCCAGTGGGAGAGGCCAGGACACTGCT 4560
DB |||||
4501 CATACTGTGTGTGGGTGTGCTGCTGCACTCCAGTGGGAGAGGCCAGGACACTGCT 4560
QY 4561 CAGCATGGTACAGTGACAGGACAGCCCCCATCATCAAGAAATTAATCTGGTCCCAATGTC 4620

Db 4561 CAGCATGGTACAGTCACAGGACAGCCCATCATCAAGAAATATCTGGTCCCAATGTC 4620
QY 4621 AATAGTTTGAGCATTTGAGAGACCCCTAGCCCTTCACTTTAAGTTTTCTGGCGTTCTCGATCT 4680
Db 4621 AATAGTTTGAGCATTTGAGAGACCCCTAGCCCTTCACTTTAAGTTTTCTGGCGTTCTCGATCT 4680
QY 4681 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAGAGTACTGGGAGTGATCAACTAGAGCCAG 4740
Db 4681 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAGAGTACTGGGAGTGAATCAACTAGAGCCAG 4740
QY 4741 GAATATATTTGGGACGCGTTTGGTGTCTGCTCAAAACCTTCTCTTCTGTCTGCGCAAG 4800
Db 4741 GAATATATTTGGGACGCGTTTGGTGTCTGCTCAAAACCTTCTCTTCTGTCTGCGCAAG 4800
QY 4801 CTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTCTATAAAATACAAGGAAT 4860
Db 4861 GTGAGCACACTGAAGACATTTTAAAGAGGCTCATTTGTCTCAGCAGAAATTTTCAGTGTAC 4920
Db 4861 GTGAGCACACTGAAGACATTTTAAAGAGGCTCATTTGTCTCAGCAGAAATTTTCAGTGTAC 4920
QY 4921 TAGTGGCATTTATAGAAAGAGAGTGCATCACTGAAGGCATGCTCACATAATATTCCTGA 4980
Db 4921 TAGTGGCATTTATAGAAAGAGAGTGCATCACTGAAGGCATGCTCACATAATATTCCTGA 4980
QY 4981 GCCCTGGTGGCGTTATCTAGGGCAAGGATTCACCTGTGTTGGAGTGGCGCCATCC 5040
Db 4981 GCCCTGGTGGCGTTATCTAGGGCAAGGATTCACCTGTGTTGGAGTGGCGCCATCC 5040
QY 5041 TCACGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTGTAATAGAGGATCTTG 5100
Db 5041 TCACGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTGTAATAGAGGATCTTG 5100
QY 5101 CTGCTTAAACAGTTGAAAGACCTGTATGGGAGCCGTAAATGACAAAGCAATGATGG 5160
Db 5101 CTGCTTAAACAGTTGAAAGACCTGTATGGGAGCCGTAAATGACAAAGCAATGATGG 5160
QY 5161 GAACATGAATCGTCTTAGGGAAGCATCTGTCAAAAGTGTCTCTGGTTAAAAACAAGTGCC 5220
Db 5161 GAACATGAATCGTCTTAGGGAAGCATCTGTCAAAAGTGTCTCTGGTTAAAAACAAGTGCC 5220
QY 5221 TCCTCTCTCAGTGTCACTGATGTGTGCTTGAATCTTTCGAAAACTGGGTGTATGAG 5280
Db 5221 TCCTCTCTCAGTGTCACTGATGTGTGCTTGAATCTTTCGAAAACTGGGTGTATGAG 5280
QY 5281 ACCCAGATGAATTTGCCCACACGATTGATTTGAGTCTTCTTCCCTGCTCTTCAGCCA 5340
Db 5281 ACCCAGATGAATTTGCCCACACGATTGATTTGAGTCTTCTTCCCTGCTCTTCAGCCA 5340
QY 5341 GTGCCAGTTCCTTTCTGATCATGTGATTTGACGTGAGAACTGTAGTCTGTATATCAAAATC 5400
Db 5341 GTGCCAGTTCCTTTCTGATCATGTGATTTGACGTGAGAACTGTAGTCTGTATATCAAAATC 5400
QY 5401 TTTAGAACTTTTGGAGTTTCTCTGGGACACAGAAACCCAGCACTTAGCATATCAAAAT 5460
Db 5401 TTTAGAACTTTTGGAGTTTCTCTGGGACACAGAAACCCAGCACTTAGCATATCAAAAT 5460
QY 5461 CTAATGCTCTTAATGGCATCATAAAGAGGCTTTTAAACACAGACTCCAGTTAGCTAAGTG 5520
Db 5461 CTAATGCTCTTAATGGCATCATAAAGAGGCTTTTAAACACAGACTCCAGTTAGCTAAGTG 5520
QY 5521 GTTTCCTGTAGTCCCGGTACTGTGTGAGGGGCCCTGTGAGATGCCCGGTTCCCTGAAAG 5580
Db 5521 GTTTCCTGTAGTCCCGGTACTGTGTGAGGGGCCCTGTGAGATGCCCGGTTCCCTGAAAG 5580
QY 5581 AATAGAAAGGCGAGTTACCGGTAGTGTGTGAAAAACATGGGCTAGATCATCAGGCGAG 5640
Db 5581 AATAGAAAGGCGAGTTACCGGTAGTGTGTGAAAAACATGGGCTAGATCATCAGGCGAG 5640
QY 5641 GACAGAACTCCCTGGCTGTGGGTGGGAGCACCCAGCTTGGCGTTGAGTTCTGGTTCTACC 5700

Db 5641 GACAGAACTCCCTGGCTGTGGGTGGGAGCACCCAGCTTGGCGTTGAGTTCTGGTTCTACC 5700
QY 5701 ACTGGTTGTTTTGTGACCAATTTATGAGTGTGCTTAAACCTTCTTCTTGTGCTACTATTTCCCTG 5760
Db 5701 ACTGGTTGTTTTGTGACCAATTTATGAGTGTGCTTAAACCTTCTTGTGCTACTATTTCCCTG 5760
QY 5761 TTTGCARAAATGGTTCAATTTGACCCCTCTTCCACCTCCCAAGGACAAATTTTCAACAGCCTA 5820
Db 5761 TTTGCARAAATGGTTCAATTTGACCCCTCTTCCACCTCCCAAGGACAAATTTTCAACAGCCTA 5820
QY 5821 TTTGTAAAAAAGATCACAGTCTCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGCTTGA 5880
Db 5821 TTTGTAAAAAAGATCACAGTCTCTTTAAAAAATATAACTGTAAAGTCAGAGGTGATGCTTGA 5880
QY 5881 AAGACAGAAACACAGGTAGTGTGAAATGTCATGCTCTTGTCTTAAAGAAAAAGCAAT 5940
Db 5881 AAGACAGAAACACAGGTAGTGTGAAATGTCATGCTCTTGTCTTAAAGAAAAAGCAAT 5940
QY 5941 TCATAGCTTTTTTGGATATGACGCAACATACCAATAATCTTGACACATAGTTGGGAGTCGG 6000
Db 5941 TCATAGCTTTTTTGGATATGACGCAACATACCAATAATCTTGACACATAGTTGGGAGTCGG 6000
QY 6001 AAATGCAACAAACGCGCAGTTTATAAAACCCAGCTAGTTTGGGTATGATTTGTAAGAAAAAAA 6060
Db 6001 AAATGCAACAAACGCGCAGTTTATAAAACCCAGCTAGTTTGGGTATGATTTGTAAGAAAAAAA 6060
QY 6061 AGCTGGCCATTTCTGATTTTGGGAAATGATTTTCTTAAACCTTATATATCTTTAGTAGTCT 6120
Db 6061 AGCTGGCCATTTCTGATTTTGGGAAATGATTTTCTTAAACCTTATATATCTTTAGTAGTCT 6120
QY 6121 AGATTTATCATATTTGCTATCATCTCTGCTTTTAAAGCTTAAAGAGATCAAGTAAT 6180
Db 6121 AGATTTATCATATTTGCTATCATCTCTGCTTTTAAAGCTTAAAGAGATCAAGTAAT 6180
QY 6181 TTTTCTTCTTCTTTTAGACACTATATAGATCATCAAGGCTGTCTGTCTTACAGGTGGAT 6240
Db 6181 TTTTCTTCTTCTTTTAGACACTATATAGATCATCAAGGCTGTCTGTCTTACAGGTGGAT 6240
QY 6241 AGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAAACTTCACTGTGTTGG 6300
Db 6241 AGTGATATGATCTACAGTGAGGGGACATTTATTTAAAACTTAAAACTTCACTGTGTTGG 6300
QY 6301 GGGTGGTATTTTAAACGCGCAGCCTCTGATTTGCTTTTGGAGGCTGGTGTGTTGAA 6360
Db 6301 GGGTGGTATTTTAAACGCGCAGCCTCTGATTTGCTTTTGGAGGCTGGTGTGTTGAA 6360
QY 6361 GTTCTGTCTCTCTTCCAGTGGACCTTAACTTCTCTGATGCACTGAGACATGTCCT 6420
Db 6361 GTTCTGTCTCTCTTCCAGTGGACCTTAACTTCTCTGATGCACTGAGACATGTCCT 6420
QY 6421 ATTGTCTGCGAGAAACTAAAGCCAAACACATGTCATCTGGGACAGGTTTTCATTTGTCTAG 6480
Db 6421 ATTGTCTGCGAGAAACTAAAGCCAAACACATGTCATCTGGGACAGGTTTTCATTTGTCTAG 6480
QY 6481 ATCTCTTTTGGCCCACTGAGTGTGTTGTGACAAATACAGCTGCTTTTCCAAAACTTTGCTA 6540
Db 6481 ATCTCTTTTGGCCCACTGAGTGTGTTGTGACAAATACAGCTGCTTTTCCAAAACTTTGCTA 6540
QY 6541 AAATTTGACAGACTTTTCTAGGTGCTTCCCAATGCCAGACTTTCTTTTCTGTTGAAGAT 6600
Db 6541 AAATTTGACAGACTTTTCTAGGTGCTTCCCAATGCCAGACTTTCTTTTCTGTTGAAGAT 6600
QY 6601 TAAGTTGTGCTTGTGCGCTCTTAGTGGTCAGTGTGTTAATCTAACTTAAACCGGCTTAT 6660
Db 6601 TAAGTTGTGCTTGTGCGCTCTTAGTGGTCAGTGTGTTAATCTAACTTAAACCGGCTTAT 6660
QY 6661 TTTTCCCTGGTGTGGGAGTGTGACCGTTTGTAAATTTGAGTGTGCTTCTAAATTTTTC 6720
Db 6661 TTTTCCCTGGTGTGGGAGTGTGACCGTTTGTAAATTTGAGTGTGCTTCTAAATTTTTC 6720
QY 6721 TGAAGAAGATAATTTTCCCGCCAGTATGTATGTCACCTTCCAGTTTGGCAGATCTGCG 6780
Db 6721 TGAAGAAGATAATTTTCCCGCCAGTATGTATGTCACCTTCCAGTTTGGCAGATCTGCG 6780

QY 3354 CAGGTGACGCGCGCAGAGCATCTCTCTCTTACAAATCTGATGGTTTGTGGCTTC 3413
Db 2881 CAGGTGACGCGCGCAGAGCATCTCTCTCTTACAAATCTGATGGTTTGTGGCTTC 2940
QY 3414 CTATTTCCAAATTAAGAAGTCTAATTGAGCACCGCAAGTGCACACCAAAAAAATCTGCT 3473
Db 2941 CTATTTCCAAATTAAGAAGTCTAATTGAGCACCGCAAGTGCACACCAAAAAAATCTGCT 3000
QY 3474 TTCCGTTACAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCTCCGAGGAG 3533
Db 3001 TTCCGTTACAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCTCCGAGGAG 3060
QY 3534 GACTTCTCGCAGTTGTTCACTTGAGACCAAAATCTCACCTTGAAACGGGGAAGAAGCTT 3593
Db 3061 GACTTCTCGCAGTTGTTCACTTGAGACCAAAATCTCACCTTGAAACGGGGAAGAAGCTT 3120
QY 3594 GTCCAGATGCATCCCTCAGCTCGATCCGTTTCCACACTTCCAGGCTTTGGCAGCTGCTACC 3653
Db 3121 GTCCAGATGCATCCCTCAGCTCGATCCGTTTCCACACTTCCAGGCTTTGGCAGCTGCTACC 3180
QY 3654 AAAGGAAAAATGTCATTTGCAAGAAGTGAAGGAATCGGGCAAGAAGGGAGCACCGAC 3713
Db 3181 AAAGGAAAAATGTCATTTGCAAGAAGTGAAGGAATCGGGCAAGAAGGGAGCACCGAC 3240
QY 3714 AACGAGGATTCAGTTCGAGAACGAGCTTGAGAGAAAACAAATAAGGGGCAAGTTGTCCAGGC 3773
Db 3241 AACGAGGATTCAGTTCGAGAACGAGCTTGAGAGAAAACAAATAAGGGGCAAGTTGTCCAGGC 3300
QY 3774 CTCTCGAAGAGAAAGAGAAAGTGCAACACTCCACGCGGAAGCCCTCCGTTGACGCGG 3833
Db 3301 CTCTCGAAGAGAAAGAGAAAGTGCAACACTCCACGCGGAAGCCCTCCGTTGACGCGG 3360
QY 3834 GATCCCAAGTTACCCAGTAGCAGGAGAACCCACTCACTGCTCCGAGTGGCGCAAGCT 3893
Db 3361 GATCCCAAGTTACCCAGTAGCAGGAGAACCCACTCACTGCTCCGAGTGGCGCAAGCT 3420
QY 3894 TTCAGAACCTACCAACAGCTGCTTGTGACCTCAGGGTCCACAAGAGGACCGGAGGGCC 3953
Db 3421 TTCAGAACCTACCAACAGCTGCTTGTGACCTCAGGGTCCACAAGAGGACCGGAGGGCC 3480
QY 3954 GCGCGGAGTCCGCCACCATGTCGTGGACGGGAGCAGCGCGGAGCGTGTCTCTGAC 4013
Db 3481 GCGCGGAGTCCGCCACCATGTCGTGGACGGGAGCAGCGCGGAGCGTGTCTCTGAC 3540
QY 4014 CTCCGCGCCCTCTGGATGAATAATGGAGCCGTGGATCGAGGGGAAGTGGTTCTGAAGAC 4073
Db 3541 CTCCGCGCCCTCTGGATGAATAATGGAGCCGTGGATCGAGGGGAAGTGGTTCTGAAGAC 3600
QY 4074 GGATCTGAGGATGGGCTTCCGAGGAATCCATCTGGGTAAAGCTGCCGTCTCCGTCCTC 4133
Db 3601 GGATCTGAGGATGGGCTTCCGAGGAATCCATCTGGGTAAAGCTGCCGTCTCCGTCCTC 3660
QY 4134 GTGCTGTTCCGCTGTGTCTGTCTCCCGTCTCCCGCTCTCTATTTCCCATCTCCAG 4193
Db 3661 GTGCTGTTCCGCTGTGTCTGTCTCCCGTCTCCCGCTCTCTATTTCCCATCTCCAG 3720
QY 4194 ACAACGCTGGCAGGAATGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTGGTA 4253
Db 3721 ACAACGCTGGCAGGAATGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTGGTA 3780
QY 4254 TCACCTCTGTGAAGTCATTTAACTCTCAGGGCTTAATTTCTCATTTCTGTGAATAACA 4313
Db 3781 TCACCTCTGTGAAGTCATTTAACTCTCAGGGCTTAATTTCTCATTTCTGTGAATAACA 3840
QY 4314 GGGTTGAGTTAAGAGTCTCTTGTCTGAAAATAATATAATATTTTAAACGTTATC 4373
Db 3841 GGGTTGAGTTAAGAGTCTCTTGTCTGAAAATAATATAATATTTTAAACGTTATC 3900
QY 4374 GTTTTGCTCAGAAAAACACTTTTAAAAAATAAATTAATTTGTGCATCCAGGCCAAATGAC 4433
Db 3901 GTTTTGCTCAGAAAAACACTTTTAAAAAATAAATTAATTTGTGCATCCAGGCCAAATGAC 3960

QY 4434 TGCTTCTTAACTGGGCGGATTTGTTCCCAATCAGTATCTGCAATGCTCGAGGCAATTT 4493
Db 3961 TGCTTCTTAACTGGGCGGATTTGTTCCCAATCAGTATCTGCAATGCTCGAGGCAATTT 4020
QY 4494 TGGTTGTCTACTGTGTGTGGGTGTCCTGCTGGCATCCAGTGGGAGAGGCCAGGGA 4553
Db 4021 TGGTTGTCTACTGTGTGTGGGTGTCCTGCTGGCATCCAGTGGGAGAGGCCAGGGA 4080
QY 4554 CACTGCTCAGCATGGTACAGTGCAGACAGCCCCCATCATCAAGAAATATCTGGTCCC 4613
Db 4081 CACTGCTCAGCATGGTACAGTGCAGACAGCCCCCATCATCAAGAAATATCTGGTCCC 4140
QY 4614 AAATGCTCAATAGTTGAGCATTTGAGAGACCTAGCTTCACTTAAGTTTCTTCGCGTTC 4673
Db 4141 AAATGCTCAATAGTTGAGCATTTGAGAGACCTAGCTTCACTTAAGTTTCTTCGCGTTC 4200
QY 4674 CTGATCTTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTTGGGAG --- TGATCAA 4730
Db 4201 CTGATCTTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTTGGGAGTGAATCAA 4260
QY 4731 CTAGAGCAGGAATATATATTTGGGAGCGCTTGTGCTGTCCAAAACTTGTCTCTTCT 4790
Db 4261 CTAGAGCAGGAATATATATTTGGGAGCGCTTGTGCTGTCCAAAACTTGTCTCTTCT 4320
QY 4791 GTCTGGCAAGCTAGTATCCATTTATAGGTACCTCAGGAACCCCAATGATTTGTCTAATA 4850
Db 4321 GTCTGGCAAGCTAGTATCCATTTATAGGTACCTCAGGAACCCCAATGATTTGTCTAATA 4380
QY 4851 TACAAGGAATGTAGCACACTGAACACATTTTAAAGAGGCTCATTTGTCTCAGCAGAAAT 4910
Db 4381 TACAAGGAATGTAGCACACTGAACACATTTTAAAGAGGCTCATTTGTCTCAGCAGAAAT 4440
QY 4911 TTCAGTGTACTAGTGGCATTTATAGAAAGAGAGGTGATCACTGAAGGATGCTCACATA 4970
Db 4441 TTCAGTGTACTAGTGGCATTTATAGAAAGAGAGGTGATCACTGAAGGATGCTCACATA 4500
QY 4971 ATATTTCTGAGCCCTGGTGGGCTTATCTAGGGCAAGGATTTCCACCTGTGTTCGAGTT 5030
Db 4501 ATATTTCTGAGCCCTGGTGGGCTTATCTAGGGCAAGGATTTCCACCTGTGTTCGAGTT 4560
QY 5031 GCGCCCATCTCCTCAGTGTAGCCAGAGCTTCTCCTATCAGAGTTTGTATTTTGTGTAATA 5090
Db 4561 GCGCCCATCTCCTCAGTGTAGCCAGAGCTTCTCCTATCAGAGTTTGTATTTTGTGTAATA 4620
QY 5091 GAGGATCTTGTCTGTAAAAACAGTTGAAAAAGACCTGATGGGAGCGCGTAATTCACAG 5150
Db 4621 GAGGATCTTGTCTGTAAAAACAGTTGAAAAAGACCTGATGGGAGCGCGTAATTCACAG 4680
QY 5151 CGAATGATGGGAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAAGTGGTCTCGTTAA 5210
Db 4681 CGAATGATGGGAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAAGTGGTCTCGTTAA 4740
QY 5211 AACAAAGTCCCTCCTCCTCAGTGTCACTTGAATGTGTCTGTAATTTCTCGAAAACTG 5270
Db 4741 AACAAAGTCCCTCCTCCTCAGTGTCACTTGAATGTGTCTGTAATTTCTCGAAAACTG 4800
QY 5271 GGTGTATGAGACCCAGATGAATTTGGCCACACAGATTTGGAATCTTCTTCCTCACCTGC 5330
Db 4801 GGTGTATGAGACCCAGATGAATTTGGCCACACAGATTTGGAATCTTCTTCCTCACCTGC 4860
QY 5331 TCTTCAGCCAGTCCAGTTCCTTTTCTGATCATGTGATTTGAGTGTGAGTGTGTGTCTGT 5390
Db 4861 TCTTCAGCCAGTCCAGTTCCTTTTCTGATCATGTGATTTGAGTGTGAGTGTGTGTGTGT 4920
QY 5391 ATATCAAAATCTTTAGAAATGTTTTGAGTTTCTGGGACACAGGAAACCCAGCACTTAGCA 5450
Db 4921 ATATCAAAATCTTTAGAAATGTTTTGAGTTTCTGGGACACAGGAAACCCAGCACTTAGCA 4980
QY 5451 TACTCAAAATCTTAATGTCTTAATGTCATATAAAAGAGGCTTTTAAACACAGACTCCAGT 5510
Db 4981 TACTCAAAATCTTAATGTCTTAATGTCATATAAAAGAGGCTTTTAAACACAGACTCCAGT 5040
QY 5511 TAGCTAAGTGGTTTCTGCTAGTGGCGGTACTGTTGACGGGGCCCTGTGTGAGTGCCTCCAGT 5570

Db 5041 |||||TAGCTAAGTGGTTCTCTCTAGTCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGT 5100
QY 5571 |||||TCCCTCGAAGAAATGAAAGCCAGTTACCGGTAGGTGGTGGGAAACATCGGCTAGAT 5630
Db 5101 |||||TCCCTCGAAGAAATGAAAGCCAGTTACCGGTAGGTGGTGGGAAACATCGGCTAGAT 5160
QY 5631 |||||CATCAGGCAGGACAGAAATGCTGGCTGTGGGTGGGAGCACCCAGCTGTGGCGTTGAGTTTC 5690
Db 5161 |||||CATCAGGCAGGACAGAAATGCTGGCTGTGGGTGGGAGCACCCAGCTGTGGCGTTGAGTTTC 5220
QY 5691 |||||TGGTTCTACCACTGGCTGTTTGTGACCAATATGAGTTGCTTAAACCTTCTTCTGCTAC 5750
Db 5221 |||||TGGTTCTACCACTGGCTGTTTGTGACCAATATGAGTTGCTTAAACCTTCTTCTGCTAC 5280
QY 5751 |||||TATTTCCCTGTTTGGCAAAATGGTTCAATGACCCCTGCTTCCACCTCCCAAGGACAAATTT 5810
Db 5281 |||||TATTTCCCTGTTTGGCAAAATGGTTCAATGACCCCTGCTTCCACCTCCCAAGGACAAATTT 5340
QY 5811 |||||CAACAGCCTATTTGTAAAAAGATCACAGTCCCTTTAAAAAATATAAATGTAAGTCAAGG 5870
Db 5341 |||||CAACAGCCTATTTGTAAAAAGATCACAGTCCCTTTAAAAAATATAAATGTAAGTCAAGG 5400
QY 5871 |||||TGATGCTTGAAGAGCAGGAAACAGGTAGATGTGGAAATGTCAATGCTTGTCTTGTAAAG 5930
Db 5401 |||||TGATGCTTGAAGAGCAGGAAACAGGTAGATGTGGAAATGTCAATGCTTGTCTTGTAAAG 5460
QY 5931 |||||AAAAAGCATTTCATAGCTTTTGGATATGACGCAACATACCATAAATCTGTGACACATGTT 5990
Db 5461 |||||AAAAAGCATTTCATAGCTTTTGGATATGACGCAACATACCATAAATCTGTGACACATGTT 5520
QY 5991 |||||TGGGAGTCGGAAATTCGAAACAGCCCGAGTTATAAACCCAGCTAGTTTGGGTATGATTCG 6050
Db 5521 |||||TGGGAGTCGGAAATTCGAAACAGCCCGAGTTATAAACCCAGCTAGTTTGGGTATGATTCG 5580
QY 6051 |||||AAGAAAAAAGCTGGCCATCTGTATTTGGGAAATGATTTTCTTAAACTTATATATC 6110
Db 5581 |||||AAGAAAAAAGCTGGCCATCTGTATTTGGGAAATGATTTTCTTAAACTTATATATC 5640
QY 6111 |||||TTAGTAGTCTAGATTTATCATATTTGACTATCATCTGGCTTTTAAAGACTTAAGAAGA 6170
Db 5641 |||||TTAGTAGTCTAGATTTATCATATTTGACTATCATCTGGCTTTTAAAGACTTAAGAAGA 5700
QY 6171 |||||TCAAGTAAATTTTCTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTT 6230
Db 5701 |||||TCAAGTAAATTTTCTTCTTTAGACACTATATAGATCATCAAGGGTGTCTGTCTT 5760
QY 6231 |||||ACAGTGGATAGTATGATCTACAGTGGGGACATTTATTTAAACCTTAAACATTCAC 6290
Db 5761 |||||ACAGTGGATAGTATGATCTACAGTGGGGACATTTATTTAAACCTTAAACATTCAC 5820
QY 6291 |||||TGTGTTTGGGGTGGTATTTTAAAGGACAGCCCTGATTTGCTTTTGGAGGGCTGGTG 6350
Db 5821 |||||TGTGTTTGGGGTGGTATTTTAAAGGACAGCCCTGATTTGCTTTTGGAGGGCTGGTG 5880
QY 6351 |||||TGTGTTTGAAGTCTGTCTCTCTCCAGTGGACTCTAACTTCTCTGTGACAGCTGAGAC 6410
Db 5881 |||||TGTGTTTGAAGTCTGTCTCTCTCCAGTGGACTCTAACTTCTCTGTGACAGCTGAGAC 5940
QY 6411 |||||ACATGTCCTATGTCCTCGAAGAACTAAAGCCAAACACTGTCATCTGGGGACAGTTTTC 6470
Db 5941 |||||ACATGTCCTATGTCCTCGAAGAACTAAAGCCAAACACTGTCATTTGGGGACAGTTTTC 6000
QY 6471 |||||CATTTGTGAGATCTCTTTCGCCACATGAGTGTGTGGACAAATACAGCTCTCTTCCAA 6530
Db 6001 |||||CATTTGTGAGATCTCTTTCGCCACATGAGTGTGTGGACAAATACAGCTCTCTTCCAA 6060
QY 6531 |||||AACTTTGTCTAAATTTTGACAGACTTCTCTAGTGTCTTGGCCCAATGCGACACTTCTTTTC 6590
Db 6061 |||||AACTTTGTCTAAATTTTGACAGACTTCTCTAGTGTCTTGGCCCAATGCGACACTTCTTTTC 6120
QY 6591 |||||TGTGAAAGATTAAGTTGTGCTTCTGCTCCCTCTAGTGGTCAAGTTGTTTAACTCTAAACCTTA 6650

Db 6121 |||||TGTGGAAGATTAAGTTGTGCTTGTCTAGTGGTCAAGTTGTTTAACTCTAAACCTTA 6180
QY 6651 |||||AAGCGCTTATTTTCCCTGGTGGTGGGAAAGTTGACGGTTGCTTAATTTGCTCATTTTTC 6710
Db 6181 |||||AAGCGCTTATTTTCCCTGGTGGTGGGAAAGTTGACGGTTGCTTAATTTGCTCATTTTTC 6240
QY 6711 |||||TAAATTTATTTCTGAAGAAGATAAATTTTCCCGCCAGTATGTATGTCACCTTCTAGTTTGGC 6770
Db 6241 |||||TAAATTTATTTCTGAAGAAGATAAATTTTCCCGCCAGTATGTATGTCACCTTCTAGTTTGGC 6300
QY 6771 |||||AGATCTGCTGCTCAGAGACACTGAGAACCCGGAAGCTGCGGGCAATTCAGTCTATGA 6830
Db 6301 |||||AGATCTGCTGCTCAGAGACACTGAGAACCCGGAAGCTGCGGGCAATTCAGTCTATGA 6360
QY 6831 |||||AATGATCTTCTTGTGATTAAGGCAAAACGAAGAACTGAATGTTTAAATAGTGTACTCTGCT 6890
Db 6361 |||||AATGATCTTCTTGTGATTAAGGCAAAACGAAGAACTGAATGTTTAAATAGTGTACTCTGCT 6420
QY 6891 |||||GTACCCAGAAAAAACAACCAAAATCATGTTTATAACACTCTAAACCTTCAAAACAACCTC 6950
Db 6421 |||||GTACCCAGAAAAAACAACCAAAATCATGTTTATAACACTCTAAACCTTCAAAACAACCTC 6480
QY 6951 |||||CAACAGCATTTGGTGTGCTAGCGGTTTGTCTTAAACCGATGTTATATAAAGAAAT 7010
Db 6481 |||||CAACAGCATTTGGTGTGCTAGCGGTTTGTCTTAAACCGATGTTATATAAAGAAAT 6540
QY 7011 |||||TTTTCATGCTTTCCAAAAATGTTTATGTCAGAAATATTTAAAGTCAAGCTTATTCAG 7070
Db 6541 |||||TTTTCATGCTTTCCAAAAATGTTTATGTCAGAAATATTTAAAGTCAAGCTTATTCAG 6600
QY 7071 |||||GTACTTCACTACCTTCTTATATAAATATTTTGTGTTTCTTAAAGATAAATAATGATGA 7130
Db 6601 |||||GTACTTCACTACCTTCTTATATAAATATTTTGTGTTTCTTAAAGATAAATAATGATGA 6660
QY 7131 |||||TGGAGAAAAATAAACAATCTTACATCTTCAAGAGAGTGTAGTTATTTGTTGGAAGTTTTC 7190
Db 6661 |||||TGGAGAAAAATAAACAATCTTACATCTTCAAGAGAGTGTAGTTATTTGTTGGAAGTTTTC 6720
QY 7191 |||||CCGTTCAAAATTAATACCTCAATATTTCACTCTCAGAACGCGATACAGGTAAGAACTTTTAT 7250
Db 6721 |||||CCGTTCAAAATTAATACCTCAATATTTCACTCTCAGAACGCGATACAGGTAAGAACTTTTAT 6780
QY 7251 |||||TTTTTAAACCATGATTTAGTTAAATATGATGATTTCTAAATTTTGTGTTGTTGTTCA 7310
Db 6781 |||||TTTTTAAACCATGATTTAGTTAAATATGATGATTTCTAAATTTTGTGTTGTTGTTCA 6840
QY 7311 |||||GATATCTGCCAGATCTTGGACTAGCTTAAAGATAAATATGATGATGTTGATTCAGT 7370
Db 6841 |||||GATATCTGCCAGATCTTGGACTAGCTTAAAGATAAATATGATGATGTTGATTCAGT 6900
QY 7371 |||||GGTTATTTTATTTCTTTTAGTGCCATTTGAACTGAGCCATTTGTTCTTATTTGTCAGTTCA 7430
Db 6901 |||||GGTTATTTTATTTCTTTTAGTGCCATTTGAACTGAGCCATTTGTTCTTATTTGTCAGTTCA 6960
QY 7431 |||||TTTTCTTTTCTTTCTTTTGTGAGACGGAGTCTTGCTCTGTCACTCGGCTGGAG 7490
Db 6961 |||||TTTTCTTTTCTTTCTTTTGTGAGACGGAGTCTTGCTCTGTCACTCGGCTGGAG 7020
QY 7491 |||||TGCAGTGTGCAATTTCCGCTCACTGACGCTTCCACCTCCCTGGTTCAAGCAATATCTCCT 7550
Db 7021 |||||TGCAGTGTGCAATTTCCGCTCACTGACGCTTCCACCTCCCTGGTTCAAGCAATATCTCCT 7080
QY 7551 |||||GCCTCAGCTCCCGCTAGTGGGATTAACAGTACTGCGCACACACCCGGCTAATTTCT 7610
Db 7081 |||||GCCTCAGCTCCCGCTAGTGGGATTAACAGTACTGCGCACACACCCGGCTAATTTCT 7140
QY 7611 |||||GTATTTTGTAGATAGATGGGTTTCCATGCTGCGCAGGCTGGTTTCCAACTCTGACC 7670
Db 7141 |||||GTATTTTGTAGATAGATGGGTTTCCATGCTGCGCAGGCTGGTTTCCAACTCTGACC 7200
QY 7671 |||||TCAAGTGTATCCGCTCACCTTGGCTCCCATAGTGTGGCTCCCATAGTGTGGGATTAC 7730
Db 7201 |||||TCAAGTGTATCCGCTCACCTTGGCTCCCATAGTGTGGCTCCCATAGTGTGGGATTAC 7260

QY	7731	AGGCGTAGCCACCGCGCCGCAAAAGTTCAATTTGTTAGTTTATGACTGCTATGTCCT	7790
Db	7261	AGGCGTAGCCACCGCGCCGCAAAAGTTCAATTTGTTAGTTTATGACTGCTATGTCCT	7320
QY	7791	GACTCTTATCTTATTAATAAGCTACAGTATTTTAAATGCTGCATCTTATGCTTTATGAT	7850
Db	7321	GACTCTTATCTTATTAATAAGCTACAGTATTTTAAATGCTGCATCTTATGCTTTATGAT	7380
QY	7851	TGAGAAATGAATAGAAATCTATTTAGTAGTCTTTGAGATTTGTAAGAGGAGCTATGACATCA	7910
Db	7381	TGAGAAATGAATAGAAATCTATTTAGTAGTCTTTGAGATTTGTAAGAGGAGCTATGACATCA	7440
QY	7911	TGATGTAGAGGCTCGGTAGATTTGAAATTTCAATCTCTTCCACTATCTGTGCACCC	7970
Db	7441	TGATGTAGAGGCTCGGTAGATTTGAAATTTCAATCTCTTCCACTATCTGTGCACCC	7500
QY	7971	TTGGCCAAAGTTATTTAAACCTTTTGTCTTTTGTCTTTTGTCTGCTGTAAGTAGAATAA	8030
Db	7501	TTGGCCAAAGTTATTTAAACCTTTTGTCTTTTGTCTTTTGTCTGCTGTAAGTAGAATAA	7560
QY	8031	TACATATTTTCCCTAGGCGTGTAGGAAGATTAAATAGTTAGAACTGTTGCTGTTAAATTT	8090
Db	7561	TACATATTTTCCCTAGGCGTGTAGGAAGATTAAATAGTTAGAACTGTTGCTGTTAAATTT	7620
QY	8091	TTCTATTGAAATAGGCAATCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTT	8150
Db	7621	TTCTATTGAAATAGGCAATCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTT	7680
QY	8151	CTGATGAGAAATCTTATGTAGTAGATCGAGAAAGCAAAAGGAGGAGAAAGGCTGTT	8210
Db	7681	CTGATGAGAAATCTTATGTAGTAGATCGAGAAAGCAAAAGGAGGAGAAAGGCTGTT	7740
QY	8211	TTCTTAAATAATAGATATTTGATCTATTTAGTGTCTTTTCACTTCTATATAAAGT	8270
Db	7741	TTCTTAAATAATAGATATTTGATCTATTTAGTGTCTTTTCACTTCTATATAAAGT	7800
QY	8271	GCCATTTCTTGCTTAGTGGAATAACCATACAAATGTGAAATTTTGTGAATATGCTGCAGC	8330
Db	7801	GCCATTTCTTGCTTAGTGGAATAACCATACAAATGTGAAATTTTGTGAATATGCTGCAGC	7860
QY	8331	CCAGAAAGATCTCTGAGGTATCACTTGGAGAGACATCAAGGAGAAACCAACCGATGT	8390
Db	7861	CCAGAAAGATCTCTGAGGTATCACTTGGAGAGACATCAAGGAGAAACCAACCGATGT	7920
QY	8391	TGCTGCTGAAGTCAAGAACGATGGTAAATAATCAGGACACTGAAGATGCACTTATTAACCGC	8450
Db	7921	TGCTGCTGAAGTCAAGAACGATGGTAAATAATCAGGACACTGAAGATGCACTTATTAACCGC	7980
QY	8451	TGACAGTGGGCAACCAAAATTTGAAAGATTTTGTGATGTCGCAAGATGTTACAGG	8510
Db	7981	TGACAGTGGGCAACCAAAATTTGAAAGATTTTGTGATGTCGCAAGATGTTACAGG	8040
QY	8511	CAGTCCACTGCAAGACGCTTAAAGAGATGCTTCTGTTTTCAGAAATGTTCTGGGAG	8570
Db	8041	CAGTCCACTGCAAGACGCTTAAAGAGATGCTTCTGTTTTCAGAAATGTTCTGGGAG	8100
QY	8571	CGCTGTCTCTCACGAGCACAAAGATATCTCAGGATTTTCCATAAAAAATGCACTGATGA	8630
Db	8101	CGCTGTCTCTCACGAGCACAAAGATATCTCAGGATTTTCCATAAAAAATGCACTGATGA	8160
QY	8631	CAGTCTGATTAAGTGAATATAAAACCTTACCTCTGCTTACCTGGACCTGTTAAAAAGAG	8690
Db	8161	CAGTCTGATTAAGTGAATATAAAACCTTACCTCTGCTTACCTGGACCTGTTAAAAAGAG	8220
QY	8691	ATCAGCAGTTGAACTCAGGCAAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCC	8750
Db	8221	ATCAGCAGTTGAACTCAGGCAAAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCC	8280
QY	8751	TCCTCCGGATGCGAGTACCAACCATTAACCTTGAAGTTAGCCCTCAAGAGAACCAACGGA	8810
Db	8281	TCCTCCGGATGCGAGTACCAACCATTAACCTTGAAGTTAGCCCTCAAGAGAACCAACGGA	8340

QY	8811	GACCGAGCTGACTGCAGATACAGCCCAAGTGTGGATTTGTACGAAAAACCTTTAAATTT	8870
Db	8341	GACCGAGCTGACTGCAGATACAGCCCAAGTGTGGATTTGTACGAAAAACCTTTAAATTT	8400
QY	8871	ATCCGTGGGGCTCTTTCAAAATTTCCCGGCAATTTCTTTGAGTAAAGTTTGATTTCCAAG	8930
Db	8401	ATCCGTGGGGCTCTTTCAAAATTTCCCGGCAATTTCTTTGAGTAAAGTTTGATTTCCAAG	8460
QY	8931	TATCACCTGTCCATTTTGTACCTTCAAGACATTTTATCCAGAAATTTTAAATGATGACCA	8990
Db	8461	TATCACCTGTCCATTTTGTACCTTCAAGACATTTTATCCAGAAATTTTAAATGATGACCA	8520
QY	8991	GAGACTGAGCATATAATCAATCTCTGAGTTTCATAAAACCTGTGCAAAACAGTCTTGTCT	9050
Db	8521	GAGACTGAGCATATAATCAATCTCTGAGTTTCATAAAACCTGTGCAAAACAGTCTTGTCT	8580
QY	9051	TAGAAGTCCAGCTACCGGATGCCCGCAGCGTGTCTGGGAAAGATGTGCTTCCCTCCCTCC	9110
Db	8581	TAGAAGTCCAGCTACCGGATGCCCGCAGCGTGTCTGGGAAAGATGTGCTTCCCTCCCTCC	8640
QY	9111	TAGTTTCTGTAAACCCAAAGCCCAAGTCTGCTTTTCCCGGCGAGTCCAAATCCCTGCCATC	9170
Db	8641	TAGTTTCTGTAAACCCAAAGCCCAAGTCTGCTTTTCCCGGCGAGTCCAAATCCCTGCCATC	8700
QY	9171	TGCGAAGGGGAGCAGAGCCCTCTCGGCCAGGCAAGCCCTCTGACTTCAGGGATAGA	9230
Db	8701	TGCGAAGGGGAGCAGAGCCCTCTCGGCCAGGCAAGCCCTCTGACTTCAGGGATAGA	8760
QY	9231	CTCTAGCACTTTTAGCCCCCAAGTAACTGAAAGTCCCACAGACACACAGCAGAAATGTGGGGT	9290
Db	8761	CTCTAGCACTTTTAGCCCCCAAGTAACTGAAAGTCCCACAGACACACAGCAGAAATGTGGGGT	8820
QY	9291	CCAGGGCCGCGCACAGGCAACAGCAATCTGAGATGTTTCTTAAACAGAGTGTTCCTCC	9350
Db	8821	CCAGGGCCGCGCACAGGCAACAGCAATCTGAGATGTTTCTTAAACAGAGTGTTCCTCC	8880
QY	9351	TGCAACCGATTAAGCAAAAGACCCGACAGCAAAATTTGAAACCTCTTCAGTAGCTCTTTC	9410
Db	8881	TGCAACCGATTAAGCAAAAGACCCGACAGCAAAATTTGAAACCTCTTCAGTAGCTCTTTC	8940
QY	9411	TCAGCCCACTCTCGGACAGTAAATCAATGTTTCCATCGACTACCCCGCCCAAGAACGA	9470
Db	8941	TCAGCCCACTCTCGGACAGTAAATCAATGTTTCCATCGACTACCCCGCCCAAGAACGA	9000
QY	9471	CAGCCCGTGGGCACTCTCGGGAAGAGACTATTTCTGTAAATCGGAGTCCAGCAATACTGC	9530
Db	9001	CAGCCCGTGGGCACTCTCGGGAAGAGACTATTTCTGTAAATCGGAGTCCAGCAATACTGC	9060
QY	9531	AGCAGAAATTTGTGAGCCCTTTCCAAAAGACTGAAGTCCAGCGTGGTTCGCCCTTGACGT	9590
Db	9061	AGCAGAAATTTGTGAGCCCTTTCCAAAAGACTGAAGTCCAGCGTGGTTCGCCCTTGACGT	9120
QY	9591	TGACAGCCCGGGGCAATTTACAGAAAGGCTATGACCTTCCAAAGTACCATATGCTCAG	9650
Db	9121	TGACAGCCCGGGGCAATTTACAGAAAGGCTATGACCTTCCAAAGTACCATATGCTCAG	9180
QY	9651	AGGCATACATCACTGTTACCGCAGGACTGTGTGTATTCGTGCGAGGCGCTGCCCTCCCA	9710
Db	9181	AGGCATACATCACTGTTACCGCAGGACTGTGTGTATTCGTGCGAGGCGCTGCCCTCCCA	9240
QY	9711	ACCAAGTTTCTGAGCTCCAGCGAGTCTGATTTCTCAAATGTGCTGACTGTTTCAAGAGC	9770
Db	9241	ACCAAGTTTCTGAGCTCCAGCGAGTCTGATTTCTCAAATGTGCTGACTGTTTCAAGAGC	9300
QY	9771	CTATGCTGGCTCCCGGCCACTTTTACATTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG	9830
Db	9301	CTATGCTGGCTCCCGGCCACTTTTACATTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG	9360
QY	9831	GACGTTAGAGGTTATGATGAGGGGCGTCTGTTTAAATGCTGCTGCTGCTGCTGCTGCTG	9890
Db	9361	GACGTTAGAGGTTATGATGAGGGGCGTCTGTTTAAATGCTGCTGCTGCTGCTGCTGCTG	9420
QY	9891	AGCTAATCCAGGCAATTTCTCAGTGGAGATGGTACCACCTCCCAAGGGTGGGGGGTAGCAGC	9950

Accession	Sequence	Length
Db	AGCTAATCCAGSCAATCTCAGTGAGATGTTACCACTCCCAAGGGTGGGGGTAGGCAGC	9480
Qy	CAGAAGTCTCTGGGGGTCCACAGAGAGAGCAATCTTAGATACGCAGTGG	10000
Db	CAGAAGTCTCTGGGGGTCCACAGAGAGAGCAATCTTAGATACGCAGTGG	9530

RESULT 3

```

RESULT 3
US-09-949-016-16004
; Sequence 16004, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16004
; LENGTH: 20023
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16004

```

Query Match	95.1%	Score 9509.2	DB 4	Length 20023	1;
Best Local Similarity	99.9%	Pred. No. 0;			
Matches 9524; Conservative	0;	Mismatches	3;	Indels	3; Gaps
Qy	474	TTCAAATTCACCTGAACCTTTGATTTTGTCTAAATAAGCAGCATTAATAAAACCAACCAAAAA	533		
Db	1	TTCAAATTCACCTGAACCTTTGATTTTGTCTAAATAAGCAGCATTAATAAAACCAACCAAAAA	60		
Qy	534	ACTCCACTATTGCCATTTGCGCCTATTGTGATTTTTTAAAAAATAAGCGTATTTTAGCATC	593		
Db	61	ACTCCACTATTGCCATTTGCGCCTATTGTGATTTTTTAAAAAATAAGCGTATTTTAGCATC	120		
Qy	594	TAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTTGTCTTGGCCAGGGAAAAACAGCGTT	653		
Db	121	TAAAGTAGGAAGGACCTCAAATAAATGAGTCTTTTGTCTTGGCCAGGGAAAAACAGCGTT	180		
Qy	654	GTCAAGATTTGATAACCTGTTTTCTAGGGGTATGTCGTGTTATTCAGTTAAAAACCTTGCCT	713		
Db	181	GTCAAGATTTGATAACCTGTTTTCTAGGGGTATGTCGTGTTATTCAGTTAAAAACCTTGCCT	240		
Qy	714	GGGACGCTAGCATTCAGTAAATACTTGTGCAATAAGCAAAATGAAACTTAAGCTTCTATGT	773		
Db	241	GGGACGCTAGCATTCAGTAAATACTTGTGCAATAAGCAAAATGAAACTTAAGCTTCTATGT	300		
Qy	774	ATGAAACCTTAAGTCACTTCACATTCCTGATTTAGCAGAGTAATTGAATATTTCTTTTCAAATG	833		
Db	301	ATGAAACCTTAAGTCACTTCACATTCCTGATTTAGCAGAGTAATTGAATATTTCTTTTCAAATG	360		
Qy	834	TGTAGCTCTATCCCCAGAACCAAGATATTGAACTGTAAAGGCCATCCTATAGTTTAA	893		
Db	361	TGTAGCTCTATCCCCAGAACCAAGATATTGAACTGTAAAGGCCATCCTATAGTTTAA	420		
Qy	894	CCAACCTCGCTTAATAGATAATAGAAGATGTGGTATGTGGCAGTGCACACTTTGAAGGTT	953		
Db	421	CCAACCTCGCTTAATAGATAATAGAAGATGTGGTATGTGGCAGTGCACACTTTGAAGGTT	480		
Qy	954	GTGACTAGAACTCGGGTCTCTGGAGTGTCTTATATATATACACAAAGCTGGTCACAGCC	1013		
Db	481	GTGACTAGAACTCGGGTCTCTGGAGTGTCTTATATATATACACAAAGCTGGTCACAGCC	540		

|||||
1621 AAAATGTAATCAGCTGAATTTAAACAATTTCTGACCTTAAACCTTGACATTTTAGATTTAAG 1680
QY
2154 TCGGTAAAGCACTGATTTAAACTGGAATTTAACTGGATGAAATTTCTGATTTAATTAAGTGT 2213
Db
1681 TCGGTAAAGCACTGATTTAAACTGGAATTTAACTGGATGAAATTTCTGATTTAATTAAGTGT 1740
QY
2214 ACTGACTGGATAAAATGCCAATGATTTAACTAAAGCAACGTTTAAACAGGATGCCCTATA 2273
Db
1741 ACTGACTGGATAAAATGCCAATGATTTAACTAAAGCAACGTTTAAACAGGATGCCCTATA 1800
QY
2274 TATTAGTTTAAAGTGAAGCAATTTGAATTAGGTACCTTCTCTGCTGCGTGGAAAGACCGT 2333
Db
1801 TATTAGTTTAAAGTGAAGCAATTTGAATTAGGTACCTTCTCTGCTGCGTGGAAAGACCGT 1860
QY
2334 ATGACTCAACCCACACACGACCTTCTCTCTGCTCTGAGTGTAGTAAACCGTTTCTGTGTTTTT 2393
Db
1861 ATGACTCAACCCACACACGACCTTCTCTCTGCTCTGAGTGTAGTAAACCGTTTCTGTGTTTTT 1920
QY
2394 TTCCCTCTAGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTTGACTTCTTGCTCAAT 2453
Db
1921 TTCCCTCTAGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTTGACTTCTTGCTCAAT 1980
QY
2454 TGAACACACTCATTTCAATGGAGACAAAGAGAACTAATGCTTTTGTGCTGATTTCAATTTTGAA 2513
Db
1981 TGAACACACTCATTTCAATGGAGACAAAGAGAACTAATGCTTTTGTGCTGATTTCAATTTGAA 2040
QY
2514 TCGAGGCATTTGGGAACCCCTGTATGCTTTGTTGTGGAAGAACCAAGTGACACCATCACTG 2573
Db
2041 TCGAGGCATTTGGGAACCCCTGTATGCTTTGTTGTGGAAGAACCAAGTGACACCATCACTG 2100
QY
2574 AGCTTCTTAAAGTTTGAAGAAAGTTAGAGGACTATACACTTCTTTTGAACCTTTTATTAAT 2633
Db
2101 AGCTTCTTAAAGTTTGAAGAAAGTTAGAGGACTATACACTTCTTTTGAACCTTTTATTAAT 2160
QY
2634 AAATATTTGCTCTGTGTTTGGAAACCCAGGGCTGTTAGAGGGTGAAGTGAACAGTCTTAC 2693
Db
2161 AAATATTTGCTCTGTGTTTGGAAACCCAGGGCTGTTAGAGGGTGAAGTGAACAGTCTTAC 2220
QY
2694 AAGTGGCCCTATTCCAACTCCAGAAATTTGCCAACCGGAACCTTTGAGATTTATATGCAATCG 2753
Db
2221 AAGTGGCCCTATTCCAACTCCAGAAATTTGCCAACCGGAACCTTTGAGATTTATATGCAATCG 2280
QY
2754 AAGTGAACAGGAAACATGCCAACTCAATCCCTCTTAATGTATACATGGATGGGCGCAAGTG 2813
Db
2281 AAGTGAACAGGAAACATGCCAACTCAATCCCTCTTAATGTATACATGGATGGGCGCAAGTG 2340
QY
2814 ATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCCTTGTCAATGAAAGGGACC 2873
Db
2341 ATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCCCTTGTCAATGAAAGGGACC 2400
QY
2874 GCTGTTGTTCCATTCGAGCTACACAAGAAAAAATGTCAATCCAAATCGAGGGGTATATG 2933
Db
2401 GCTGTTGTTCCATTCGAGCTACACAAGAAAAAATGTCAATCCAAATCGAGGGGTATATG 2460
QY
2934 CCCTTGGATTTGATGTTCTGACGACAGACCTTCAACATTCAGAGACCTTAAATAACAT 2993
Db
2461 CCCTTGGATTTGATGTTCTGACGACAGACCTTCAACATTCAGAGACCTTAAATAACAT 2520
QY
2994 GTCTTAAATGCAACACCGGCTTACCCTCTGTGAACAGCAGAGTTCTTTCGGGTTCAGCAGAG 3053
Db
2521 GTCTTAAATGCAACACCGGCTTACCCTCTGTGAACAGCAGAGTTCTTTCGGGTTCAGCAGAG 2580
QY
3054 TATCTCAGTCCGCTTGATAAAAAGTCAAGTGCAGAACCAACCTCCCAAGGAAAAAAGATTGC 3113
Db
2581 TATCTCAGTCCGCTTGATAAAAAGTCAAGTGCAGAACCAACCTCCCAAGGAAAAAAGATTGC 2640
QY
3114 AAGGAAATGAAATTTAGCTGTAGGTATGTGGGCGACATTTAGAGTGCCTTTTGAATGTT 3173
Db
2641 AAGGAAATGAAATTTAGCTGTAGGTATGTGGGCGACATTTAGAGTGCCTTTTGAATGTT 2700
QY
3174 GAGATCCACATGAGAAACACAAAAGATTTCTTCACTTACGGGTGTAAACATGTGCGGGAAGA 3233
|||||

Db
2701 GAGATCCACATGATGAGAACACACAAAGATTTCTTCACTTACGGGTGTAAACATGTGCGGAAGA 2760
QY
3234 AGATTTAAAGGAGCCTTGTGTTTCTTAAAAATCACATGCGGACACATAATGGGAAATTCGGGG 3293
Db
2761 AGATTTAAAGGAGCCTTGTGTTTCTTAAAAATCACATGCGGACACATAATGGGAAATTCGGGG 2820
QY
3294 GCCAGAACAACTGACAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACAGAGTCTGTC 3353
Db
2821 GCCAGAACAACTGACAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACAGAGTCTGTC 2880
QY
3354 CAGGTGACGCGGGCGGAGAGCATCTCTCTCTCTTACAAAAATCTGATGTTTGTGGCTTC 3413
Db
2881 CAGGTGACGCGGGCGGAGAGCATCTCTCTCTCTTACAAAAATCTGATGTTTGTGGCTTC 2940
QY
3414 CTATTTCCAAATAAAGAAAGTCTTAATTTAGACACCGCAAGGTGCAACCAAAAAAACTGCT 3473
Db
2941 CTATTTCCAAATAAAGAAAGTCTAATTTAGACACCGCAAGGTGCAACCAAAAAAACTGCT 3000
QY
3474 TTTCGTACACAGCAGCGCGACAGACACTCTCCAAAGGAGGAATGCCGTCTCTCGAGGGAG 3533
Db
3001 TTTCGTACACAGCAGCGCGACAGACACTCTCCAAAGGAGGAATGCCGTCTCTCGAGGGAG 3060
QY
3534 GACTTCTCTGAGTTGTTTCAACTTGAGACCAAAATCTCACTCTGAAACCGGGGAAGAGCCT 3593
Db
3061 GACTTCTCTGAGTTGTTTCAACTTGAGACCAAAATCTCACTCTGAAACCGGGGAAGAGCCT 3120
QY
3594 GTGAGATGATCCCTCAGCTCGATCCGTTTACCACTTCCAGGCTTCCAGGCTTGGCAGCTGCTTACC 3653
Db
3121 GTGAGATGATCCCTCAGCTCGATCCGTTTACCACTTCCAGGCTTCCAGGCTTGGCAGCTGCTTACC 3180
QY
3654 AAAGGAAAGTTGCCATTTGCCAAGAGTGAAGGAATCGGGCAAGAAAGGAGGACCCGAC 3713
Db
3181 AAAGGAAAGTTGCCATTTGCCAAGAGTGAAGGAATCGGGCAAGAAAGGAGGACCCGAC 3240
QY
3714 AACGACGATTCGAGTTCCGAGAGAGAGCTTGAGAAAAACAAATAAGGGCAGTTGTGACGAC 3773
Db
3241 AACGACGATTCGAGTTCCGAGAGAGAGCTTGAGAAAAACAAATAAGGGCAGTTGTGACGAC 3300
QY
3774 CTCTCGAAGAGAAAGAGAGTGCMAACACTTCCGCGGAGAGCGCCCTCCGTGACGCGG 3833
Db
3301 CTCTCGAAGAGAAAGAGAGTGCMAACACTTCCGCGGAGAGCGCCCTCCGTGACGCGG 3360
QY
3834 GATCCCAAGTTTACCCAGTAGCAAGAGAGAGCCACTCACTGCTCCGAGTGGCGCAAGCT 3893
Db
3361 GATCCCAAGTTTACCCAGTAGCAAGAGAGAGCCACTCACTGCTCCGAGTGGCGCAAGCT 3420
QY
3894 TTCAAGACCTTACCAACAGCTGCTTTCATCTCCAGGGTCCACAAAGAGACCGGAGGGCC 3953
Db
3421 TTCAAGACCTTACCAACAGCTGCTTTCATCTCCAGGGTCCACAAAGAGACCGGAGGGCC 3480
QY
3954 GCGCGGAGTGCACCACCATGCTGTGACGCGGAGGACCGGGGACGTTCTCTCTGAC 4013
Db
3481 GCGCGGAGTGCACCACCATGCTGTGACGCGGAGGACCGGGGACGTTCTCTCTGAC 3540
QY
4014 CTGCGCGCCCTCTCGATGAAATTCGAGCCGTTGAGTCGAGGGAAGGTGTTCTGAAGAC 4073
Db
3541 CTGCGCGCCCTCTCGATGAAATTCGAGCCGTTGAGTCGAGGGAAGGTGTTCTGAAGAC 3600
QY
4074 GGATCTGAGGATGGGCTTCCCGAAGGAATTCATCTGGGTAAAGTGCCTGTCTCCGTGCC 4133
Db
3601 GGATCTGAGGATGGGCTTCCCGAAGGAATTCATCTGGGTAAAGTGCCTGTCTCCGTGCC 3660
QY
4134 GTGCTGTTCCGCTGTGTCTGTCTCCCGCTCTCCCTCTCTATTTCCATCTCCAG 4193
Db
3661 GTGCTGTTCCGCTGTGTCTGTCTCCCGCTCTCCCTCTCTATTTCCATCTCCAG 3720
QY
4194 ACAACGCTGGCCAGGAATGGGCTTTGGAGAGCGCAGAGTCAAGTCCAGGCTCTTTTGTGTA 4253
Db
3721 ACAACGCTGGCCAGGAATGGGCTTTGGAGAGCGCAGAGTCAAGTCCAGGCTCTTTTGTGTA 3780
QY
4254 TCACCTCTGTGTAAGTCAATTTAACTCTCAGGGCTTAAATTTTCTCATTTCTGTAATAACA 4313
Db
3781 TCACCTCTGTGTAAGTCAATTTAACTCTCAGGGCTTAAATTTTCTCATTTCTGTAATAACA 3840

Db	6001	 CATTTGTCAGATCTCTTTTCGCCACATAGTGTGTTGTGGACAATACAGACCTGCTTTCCAA	6060
Qy	6531	AACCTTTGCTAAATTTTGACAGACTTTCTTAGTGCTTCCCAATGCCAGACTTTCTTTTC	6590
Db	6061	AACCTTTGCTAAATTTTGACAGACTTTCTTAGTGCTTCCCAATGCCAGACTTTCTTTTC	6120
Qy	6591	TGTTGAAGATTAAGTTGTGCTTGTGCTGCCCTCTAGTGGTCAAGTTGTTTAATCCTAACCTTA	6650
Db	6121	TGTTGAAGATTAAGTTGTGCTTGTGCTGCCCTCTAGTGGTCAAGTTGTTTAATCCTAACCTTA	6180
Qy	6651	AACGGCTTAATTTTCCCTGGTGGTGGGAAGTTGACGGTTGTTAAATGGGCTCAATTTTC	6710
Db	6181	AACGGCTTAATTTTCCCTGGTGGTGGGAAGTTGACGGTTGTTAAATGGGCTCAATTTTC	6240
Qy	6711	TAAATTAATTTCTGAAGAAGATAATTTTCCCGCCAGTATGATGTGCCACCTTCAGTTGGC	6770
Db	6241	TAAATTAATTTCTGAAGAAGATAATTTTCCCGCCAGTATGATGTGCCACCTTCAGTTGGC	6300
Qy	6771	AGATCTGCTGCTCAGACACACTGAGAACCGGAAGCTGCCCGGCAATTCAGTCTATGA	6830
Db	6301	AGATCTGCTGCTCAGACACACTGAGAACCGGAAGCTGCCCGGCAATTCAGTCTATGA	6360
Qy	6831	AATGATCTTTTCTTGTGATTAAGGCAAAACGAAGAACTGAAATGTTAAATAGTGTACTCTGCT	6890
Db	6361	AATGATCTTTTCTTGTGATTAAGGCAAAACGAAGAACTGAAATGTTAAATAGTGTACTCTGCT	6420
Qy	6891	GTACCCAGAAAAAACAACAAATCATGTTATTAACACTCTAAACCTTCAACACACTC	6950
Db	6421	GTACCCAGAAAAAACAACAAATCATGTTATTAACACTCTTAAACCTTCAACACACTC	6480
Qy	6951	CAACAGCATTTGGTGTGCTGTAGCGGTTTGTCTTAACCCGATGTTATATATAAAGAAAT	7010
Db	6481	CAACAGCATTTGGTGTGCTGTAGCGGTTTGTCTTAACCCGATGTTATATATAAAGAAAT	6540
Qy	7011	TTTTTCATGCTTTCCAAAAATGTTTATGTCAAGAAATATTTAAAGTCAGCATGCTTTATTCAG	7070
Db	6541	TTTTTCATGCTTTCCAAAAATGTTTATGTCAAGAAATATTTAAAGTCAGCATGCTTTATTCAG	6600
Qy	7071	GTACTTCAGTACCTCTTATATAAATATTTTGTGTTTTTCCCTTAAAGATAAAATGATGA	7130
Db	6601	GTACTTCAGTACCTCTTATATAAATATTTTGTGTTTTTCCCTTAAAGATAAAATGATGA	6660
Qy	7131	TGGAGGAAAAATAAAACATCTTACATCTTCAAGAGAGTGATGTTATTTGTGGAAGTTTTT	7190
Db	6661	TGGAGGAAAAATAAAACATCTTACATCTTCAAGAGAGTGATGTTATTTGTGGAAGTTTTT	6720
Qy	7191	CCGTTCAAAATTAATCCTCAATTAATCATCTCAGAACGCATACAGGTAAGAACTTTTAT	7250
Db	6721	CCGTTCAAAATTAATCCTCAATTAATCATCTCAGAACGCATACAGGTAAGAACTTTTAT	6780
Qy	7251	TTTTTAAACCATGATAGTTAAATTAATGATGATTAATTAATTTTGTGTTGTTGTTCA	7310
Db	6781	TTTTTAAACCATGATAGTTAAATTAATGATGATTAATTAATTTTGTGTTGTTGTTCA	6840
Qy	7311	GATACCTGCCAGATCCTTGGACTAGCTTTAAGGATAAATATGATGATGTTGATGAGT	7370
Db	6841	GATACCTGCCAGATCCTTGGACTAGCTTTAAGGATAAATATGATGATGTTGATGAGT	6900
Qy	7371	GGTTATTTTATTTTATGTCGCAATGTAACTTGAGCCATGTTCTTATTTTGCAAGTTCA	7430
Db	6901	GGTTATTTTATTTTATGTCGCAATGTAACTTGAGCCATGTTCTTATTTTGCAAGTTCA	6960
Qy	7431	TTTCTTTTCTTTCTTTTGTGTTTGTGAGACGGAGTCTGTGTCACCTCGCGCTGAG	7490
Db	6961	TTTCTTTTCTTTCTTTTGTGTTTGTGAGACGGAGTCTGTGTCACCTCGCGCTGAG	7020
Qy	7491	TGCAGTGGTGCAATTTTGGCTCAGCTGCAGCTCCACCTCCCTGGTTCAAGCAATCTCCT	7550
Db	7021	TGCAGTGGTGCAATTTTGGCTCAGCTGCAGCTCCACCTCCCTGGTTCAAGCAATCTCCT	7080
Qy	7551	GCCTCAGCCTCCCGAGTGTGGGATTAAGGATACCTGCTGCCACACACCCGCTAAATTTCT	7610

```
QY 8691 ATCAGCAGTTGAAACTCAGGCAATAAACCTCATCTGTAGAACCAAGCGGATGTTACTCC 8750
DB 8221 ATCAGCAGTTGAAACTCAGGCAATAAACCTCATCTGTAGAACCAAGCGGATGTTACTCC 8280
QY 8751 TCCTCCGGATGGCAGTACCAACCCATACCTTGAGTTAGTCCCAAGAGCAAGCGGA 8810
DB 8281 TCCTCCGGATGGCAGTACCAACCCATACCTTGAGTTAGTCCCAAGAGCAAGCGGA 8340
QY 8811 GACCGCAGCTGACTCGACAGATACAGGCCAAGTGTGATTTGTACGAAAAACCTTTAAATTT 8870
DB 8341 GACCGCAGCTGACTCGACAGATACAGGCCAAGTGTGATTTGTACGAAAAACCTTTAAATTT 8400
QY 8871 ATCCGTGGGGCTCTTCACAATTCGCGGCAATTTCTTTGAGTAAAGTTTGAATCCAG 8930
DB 8401 ATCCGTGGGGCTCTTCACAATTCGCGGCAATTTCTTTGAGTAAAGTTTGAATCCAG 8460
QY 8931 TATCACTGTGTCATTTTGTACCTTCAAGACATTTTATCCAGAAAGTTTAAATGATGACCA 8990
DB 8461 TATCACTGTGTCATTTTGTACCTTCAAGACATTTTATCCAGAAAGTTTAAATGATGACCA 8520
QY 8991 GAGACTGGAGCTAATAATACAACTCTGACGTTTATAAAAACTGTGAAAAAGTCCCTGCT 9050
DB 8521 GAGACTGGAGCTAATAATACAACTCTGACGTTTATAAAAACTGTGAAAAAGTCCCTGCT 8580
QY 9051 TAGAAGTCGACGTACCGGATGCCCGCAGCGTTGCTGGGAAAAAGATGTGCTCCCTCCC 9110
DB 8581 TAGAAGTCGACGTACCGGATGCCCGCAGCGTTGCTGGGAAAAAGATGTGCTCCCTCCC 8640
QY 9111 TAGTTTCTGTAAACCCAGCCAAAGTCTGTTTCCCGGCGCAGTCCAAATCCCTGCCATC 9170
DB 8641 TAGTTTCTGTAAACCCAGCCAAAGTCTGTTTCCCGGCGCAGTCCAAATCCCTGCCATC 8700
QY 9171 TGGCAAGGGGAAGCAGAGCCCTCTGGGCGAGCAAGGCCCTCTGACTTCAGGGATAGA 9230
DB 8701 TGGCAAGGGGAAGCAGAGCCCTCTGGGCGAGCAAGGCCCTCTGACTTCAGGGATAGA 8760
QY 9231 CTCTAGCATTATAGCCCCAAGTAACTGAACTGCCACAGACCAAGCAAGTGTGGGGT 9290
DB 8761 CTCTAGCATTATAGCCCCAAGTAACTGAACTGCCACAGACCAAGCAAGTGTGGGGT 8820
QY 9291 CCAAGGGGGCCGACACAGGCAAGCACTGAGATGTTTCTTAAACCAAGTGTTCCTCC 9350
DB 8821 CCAAGGGGGCCGACACAGGCAAGCACTGAGATGTTTCTTAAACCAAGTGTTCCTCC 8880
QY 9351 TGCACCGGATAGACAAAAGACCCGAGACAAATTTGAAACCTTCCAGTAGCTCTTC 9410
DB 8881 TGCACCGGATAGACAAAAGACCCGAGACAAATTTGAAACCTTCCAGTAGCTCTTC 8940
QY 9411 TCAGCCACCCCTCGGCGAGCAATCAATGTTTCCATCGACTACCCCGCCCAAGAACGA 9470
DB 8941 TCAGCCACCCCTCGGCGAGCAATCAATGTTTCCATCGACTACCCCGCCCAAGAACGA 9000
QY 9471 CAGCCCGTGGGCACTCCCGGGAAGAGACTATTTCTGTAATCGGAGTCCGAGCAATCTGC 9530
DB 9001 CAGCCCGTGGGCACTCCCGGGAAGAGACTATTTCTGTAATCGGAGTCCGAGCAATCTGC 9060
QY 9531 AGCAGAAATTTGGTGGCCCTTCCAAAAGACTGAAGTCCAGCGTGGTGGCTTGAAGT 9590
DB 9061 AGCAGAAATTTGGTGGCCCTTCCAAAAGACTGAAGTCCAGCGTGGTGGCTTGAAGT 9120
QY 9591 TGACCAAGCCCGGGGCAATTTACAGAAAGAGCTATGACCTTCCCAAGTACCATATGTCAG 9650
DB 9121 TGACCAAGCCCGGGGCAATTTACAGAAAGAGCTATGACCTTCCCAAGTACCATATGTCAG 9180
QY 9651 AGGCATCAATCATCTTTACCGCAGGACTGTGTGATCCGTCCGAGCGCTGCTCCCAA 9710
DB 9181 AGGCATCAATCATCTTTACCGCAGGACTGTGTGATCCGTCCGAGCGCTGCTCCCAA 9240
QY 9711 ACCAAGTTTCTGAGCTCCAGGAGTCCGATTTCTCCAAATGCTGACTGTTTCAGAGCC 9770
DB 9241 ACCAAGTTTCTGAGCTCCAGGAGTCCGATTTCTCCAAATGCTGACTGTTTCAGAGCC 9300
```

```
QY 9771 CTATGTGGCTCCGGGCCACTTTTACACTTGTGTGCTGCTGTAGTCCAGCATCCAGCTC 9830
DB 9301 CTATGTGGCTCCGGGCCACTTTTACACTTGTGTGCTGCTGTAGTCCAGCATCCAGCTC 9360
QY 9831 GAGCTTAGAAGGATTTGCATGAGGGGCTGCTGTTTAAATGGCTGCCTACAGTGAATTA 9890
DB 9361 GAGCTTAGAAGGATTTGCATGAGGGGCTGCTGTTTAAATGGCTGCCTACAGTGAATTA 9420
QY 9891 AGCTAATCCAGGCAATTTCTCAGTGGAGATGTACCACTCCCAAGGGTGGGGGTAGGCAGC 9950
DB 9421 AGCTAATCCAGGCAATTTCTCAGTGGAGATGTACCACTCCCAAGGGTGGGGGTAGGCAGC 9480
QY 9951 CAGAAGTTCTTGGGGTTCACAGAGAGACATTTCTTAGATACGGCAGTGG 10000
DB 9481 CAGAAGTTCTTGGGGTTCACAGAGAGACATTTCTTAGATACGGCAGTGG 9530
```

RESULT 4

```
US-09-560-594-3
; Sequence 3, Application US/09560594
; Patent No. 6242590
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowser
; TITLE OF INVENTION: ANTISENSE MODULATION OF ZINC FINGER PROTEIN-217 EXPRESSION
; FILE REFERENCE: RTS-0144
; CURRENT APPLICATION NUMBER: US/09/560,594
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 3
; LENGTH: 5632
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (272)...(3418)
US-09-560-594-3
```

```
Query Match 16.4%; Score 1639.4; DB 3; Length 5632;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1640; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 2474 GACAAAGAGAACTAATGCTTTGTGCTGATTTCATATTGAATCGAGGCAATGGGAACCTG 2533
DB 1 GACAAAGAGAACTAATGCTTTGTGCTGATTTCATATTGAATCGAGGCAATGGGAACCTG 60
QY 2534 TATGCTTTGTTGTGAAAGAACCACTGACCATCCTGAGCTTCTTAAAGTTCGAAG 2593
DB 61 TATGCTTTGTTGTGAAAGAACCACTGACCATCCTGAGCTTCTTAAAGTTCGAAG 120
QY 2594 AAGTTAGAGGACTATACACTTTCTTTTGAACCTTTTATAATAATTTGCTCTGGTTTTT 2653
DB 121 AAGTTAGAGGACTATACACTTTCTTTTGAACCTTTTATAATAATTTGCTCTGGTTTTT 180
QY 2654 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACCAAGTCTTACAAGTGGCTTTATTCGAATC 2713
DB 181 GGAACCCAGGGCTGTTAGAGGGGTGAGTGACCAAGTCTTACAAGTGGCTTTATTCGAATC 240
QY 2714 CAGAAATTCGCCAACCGAACTTTGAGATTATATGCATCGAAAGTGCAGAGGAACAATGCC 2773
DB 241 CAGAAATTCGCCAACCGAACTTTGAGATTATATGCATCGAAAGTGCAGAGGAACAATGCC 300
QY 2774 AACTCAATCCCTCTTAATGTACATGATGGGCCAGAAAGTATGGCAGCTCTCTTGGCAG 2833
DB 301 AACTCAATCCCTCTTAATGTACATGATGGGCCAGAAAGTATGGCAGCTCTCTTGGCAG 360
QY 2834 TCCGATGGAGATGGAGGATGCTTGTCAATGAAGAGGACCGCTGTTGTTCCATTCGAGC 2893
DB 361 TCCGATGGAGATGGAGGATGCTTGTCAATGAAGAGGACCGCTGTTGTTCCATTCGAGC 420
QY 2894 TACACAAGAAAAAATGTGCATCCAAATCGAGGGGTATATGCCCTTGGATTCGATGTTCTG 2953
DB 421 TACACAAGAAAAAATGTGCATCCAAATCGAGGGGTATATGCCCTTGGATTCGATGTTCTG 480
```


QY 2954 CAGCCAGACCTTACACATTCAGAGACCTTAAATAACATGCTTAAATGCAACACGGCC 3013
DB 481 CAGCCAGACCTTACACATTCAGAGACCTTAAATAACATGCTTAAATGCAACACGGCC 540
QY 3014 TACCCTCTGTGAACACGAGCTTCTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 3073
DB 541 TACCCTCTGTGAACACGAGCTTCTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 600
QY 3074 AAGTCAAGTGGCAACGAAACCTCCAGGAAAGAAATTCGAAGGAAATGAATTTAGCTG 3133
DB 601 AAGTCAAGTGGCAACGAAACCTCCAGGAAAGAAATTCGAAGGAAATGAATTTAGCTG 660
QY 3134 TCAGGTATGTGGCGACAGATTTAGAGTCGCTTTGATGTTGAGATCCCATGAGAACACA 3193
DB 661 TCAGGTATGTGGCGACAGATTTAGAGTCGCTTTGATGTTGAGATCCCATGAGAACACA 720
QY 3194 CAAAGATCTCTTCACTTACGGGTGTAACATGTGCGGAAGAAATTCGAAGGACCTTGTT 3253
DB 721 CAAAGATCTCTTCACTTACGGGTGTAACATGTGCGGAAGAAATTCGAAGGACCTTGTT 780
QY 3254 TCTTAAATCATGCGGACACATATGCGAAATCGGGGGCCAGAACAACTGCGACA 3313
DB 781 TCTTAAATCATGCGGACACATATGCGAAATCGGGGGCCAGAACAACTGCGACA 840
QY 3314 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGTCTGTCAGGTGCGCGGCCGAGAG 3373
DB 841 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGTCTGTCAGGTGCGCGGCCGAGAG 900
QY 3374 CATCTCCTCTCTTCAAAATCTGCATGCTTTGCTGCTTCTTATTTCCAAATAAAGAAAG 3433
DB 901 CATCTCCTCTCTTCAAAATCTGCATGCTTTGCTGCTTCTTATTTCCAAATAAAGAAAG 960
QY 3434 TCTAATTTAGACCGCAAGGTGACACCAAAAAACTGCTTTTCGGTACGAGCGCGCA 3493
DB 961 TCTAATTTAGACCGCAAGGTGACACCAAAAAACTGCTTTTCGGTACGAGCGCGCA 1020
QY 3494 GACAGACTCTCACAGGAGGATCCGTCCTCGAGGAGGACTTCTCGAGTTGTTCAA 3553
DB 1021 GACAGACTCTCACAGGAGGATCCGTCCTCGAGGAGGACTTCTCGAGTTGTTCAA 1080
QY 3554 CTTTGAGACCAAAATCTCACCTTGAACGGGGAAGAGCTGTGAGATGATCCCTCAGCT 3613
DB 1081 CTTTGAGACCAAAATCTCACCTTGAACGGGGAAGAGCTGTGAGATGATCCCTCAGCT 1140
QY 3614 CGATCCGTTACACCTTCAGGCTTGGCAGCTGCTACCAAGGAAAGTTGCCATTG 3673
DB 1141 CGATCCGTTACACCTTCAGGCTTGGCAGCTGCTACCAAGGAAAGTTGCCATTG 1200
QY 3674 CCAAGAAGTGAAGGAATCGGGCAAGAGGGAGCACCGACACGAGTTCCGATTCGGA 3733
DB 1201 CCAAGAAGTGAAGGAATCGGGCAAGAGGGAGCACCGACACGAGTTCCGATTCGGA 1260
QY 3734 GAAGGAGCTTGGAGAAACAATAAGGCGAGTTGTGAGGCTCTCTCGCAAGAGAAAGAGAA 3793
DB 1261 GAAGGAGCTTGGAGAAACAATAAGGCGAGTTGTGAGGCTCTCTCGCAAGAGAAAGAGAA 1320
QY 3794 GTGCAAACTCCCGAGCGAAGCGCTTCGTTGAGCGGATCCCAAGTTACCCAGTAG 3853
DB 1321 GTGCAAACTCCCGAGCGAAGCGCTTCGTTGAGCGGATCCCAAGTTACCCAGTAG 1380
QY 3854 CAAGGAGAGCCACTCACTGCTCCGAGTCCGCAAGCTTTCAGAACCTTACACAGCT 3913
DB 1381 CAAGGAGAGCCACTCACTGCTCCGAGTCCGCAAGCTTTCAGAACCTTACACAGCT 1440
QY 3914 GGTCTTGTGACTCCAGGCTCCAAAGAGACCGGGCGGGGAGTCCGCCACCAT 3973
DB 1441 GGTCTTGTGACTCCAGGCTCCAAAGAGACCGGGCGGGGAGTCCGCCACCAT 1500
QY 3974 GTCTGTGAGCGGAGCGAGCGGAGCTGTTCTCTGACTCGCGCCCTCTGATGA 4033
DB 1501 GTCTGTGAGCGGAGCGAGCGGAGCTGTTCTCTGACTCGCGCCCTCTGATGA 1560
QY 4034 AAATGAGCCGTGGATCGAGGGGAAGGTGTTCTGAAAGCGGATCTGAGGATGGGCTTCC 4093

DB 1561 AAATGAGCCGTGGATCCAGGGGAAGGTGTTCTGAAGACGATCTGAGATGGGCTTCC 1620
QY 4094 CGAAGGAATCCATCTGGGTAA 4114
DB 1621 CGAAGGAATCCATCTGGGTAA 1641
RESULT 5
US-09-949-016-862
; Sequence 862, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 862
; LENGTH: 5632
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-862
Query Match 16.4%; Score 1639.4; DB 4; Length 5632;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1640; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2474 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCTTG 2533
DB 1 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCTTG 60
QY 2534 TATGCTTTGTTGTGAAAGAACCAAGTGACACCATCACTGAGCTTCTTAAAGTTTCGAAG 2593
DB 61 TATGCTTTGTTGTGAAAGAACCAAGTGACACCATCACTGAGCTTCTTAAAGTTTCGAAG 120
QY 2594 AAGTTAGAGGACTATACACTTTTGTGAACTTTTAAATAATTTGCTCTGGTTTTT 2653
DB 121 AAGTTAGAGGACTATACACTTTTGTGAACTTTTAAATAATTTGCTCTGGTTTTT 180
QY 2654 GGAACCCAGGCTGTAGAGGGTGAGTGACAAAGTCTTACAAAGTGGCCTTATTCGAACCTC 2713
DB 181 GGAACCCAGGCTGTGTAGAGGGTGAGTGACAAAGTCTTACAAAGTGGCCTTATTCGAACCTC 240
QY 2714 CAGAAATTTGCCCAACCGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 2773
DB 241 CAGAAATTTGCCCAACCGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300
QY 2774 AACTCAATCCCTTTTAATGTACATCGATGGGCGAGAGTGATGGCAGCTCTCTTGGCAG 2833
DB 301 AACTCAATCCCTTTTAATGTACATCGATGGGCGAGAGTGATGGCAGCTCTCTTGGCAG 360
QY 2834 TCCGATGAGATGGAGGATGCCCTGCTCAATGAAAGGACCGCTGTGTTCCATTCGGAGC 2893
DB 361 TCCGATGAGATGGAGGATGCCCTGCTCAATGAAAGGACCGCTGTGTTCCATTCGGAGC 420
QY 2894 TACACAGAAAAAATGTCAATCCAAATCGAGGGGTATATGCCCTTGGATTGATGTTCTG 2953
DB 421 TACACAGAAAAAATGTCAATCCAAATCGAGGGGTATATGCCCTTGGATTGATGTTCTG 480
QY 2954 CAGCCAGACCTTTCACATTCAGAGACCTTAAATAAAGATGCTTAAATGCAACACCGGCC 3013
DB 481 CAGCCAGACCTTTCACATTCAGAGACCTTAAATAAAGATGCTTAAATGCAACACCGGCC 540

QY 3014 TACCCTCTGTGAACACGACGTTCTTCGGGTGTAAGCAGATATCTCAGTCCGCTTGATAA 3073
DB |||||
DB 541 TACCCTCTGTGAACACGACGTTCTTCGGGTGTAAGCAGATATCTCAGTCCGCTTGATAA 600
QY 3074 AAGTCAAGTCGGAACAGACCTCCCAAGGAAAGAAATTCGAAGAAATCAATTTAGCTG 3133
DB |||||
DB 601 AAGTCAAGTCGGAACAGACCTCCCAAGGAAAGAAATTCGAAGAAATCAATTTAGCTG 660
QY 3134 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 3193
DB |||||
DB 661 TGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA 720
QY 3194 CAAAGATCTTTTCACTTACGCGGTGAACATGTCGCGAAGAAATTCGAAGAGCCTTGGTT 3253
DB |||||
DB 721 CAAAGATCTTTTCACTTACGCGGTGAACATGTCGCGAAGAAATTCGAAGAGCCTTGGTT 780
QY 3254 TCTTAAATAATCACATCGGACACATAATGCAATCGGGGCCAGAGCAAACTGCAGCA 3313
DB |||||
DB 781 TCTTAAATAATCACATCGGACACATAATGCAATCGGGGCCAGAGCAAACTGCAGCA 840
QY 3314 AGGCTTTGGAGAGTAGTCCAGCAACGATCAACGAGGTGCTCCAGGTGCACGCGGCCGAGAG 3373
DB |||||
DB 841 AGGCTTTGGAGAGTAGTCCAGCAACGATCAACGAGGTGCTCCAGGTGCACGCGGCCGAGAG 900
QY 3374 CATCTCCTCTCTTACAAATCTGCATGTTGTTGCTTCTTATTTCCAAATAAGAAAG 3433
DB |||||
DB 901 CATCTCCTCTCTTACAAATCTGCATGTTGTTGCTTCTTATTTCCAAATAAGAAAG 960
QY 3434 TCTAATTTGAGCAGCGGACAGTGCACACCAAAAACTGCTTTCGTTACGAGCAGCGCA 3493
DB |||||
DB 961 TCTAATTTGAGCAGCGGACAGTGCACACCAAAAACTGCTTTCGTTACGAGCAGCGCA 1020
QY 3494 GACAGACTCTCCACAAAGGAGGAATGCGCTCTCGAGGAGGACTTCTCGCAGTTGTTCAA 3553
DB |||||
DB 1021 GACAGACTCTCCACAAAGGAGGAATGCGCTCTCGAGGAGGACTTCTCGCAGTTGTTCAA 1080
QY 3554 CTTGAGACCAAAATCTCAACCTGAAACGGGGAAGAGCTGTCCAGATGCATCCCTCAGCT 3613
DB |||||
DB 1081 CTTGAGACCAAAATCTCAACCTGAAACGGGGAAGAGCTGTCCAGATGCATCCCTCAGCT 1140
QY 3614 CGATCGGTTCCACACCTTCCAGCTTCCAGCTGCTTACCAAGGAAAGTTGCCATTTG 3673
DB |||||
DB 1141 CGATCGGTTCCACACCTTCCAGCTTCCAGCTGCTTACCAAGGAAAGTTGCCATTTG 1200
QY 3674 CCAAGAGTGAAGGAATCGGGGCAAGAGGAGCACCACAGACGATTCGAGTTCCCA 3733
DB |||||
DB 1201 CCAAGAGTGAAGGAATCGGGGCAAGAGGAGCACCACAGACGATTCGAGTTCCCA 1260
QY 3734 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGAGGCTCTCCGAAGAGAAAGAGAA 3793
DB |||||
DB 1261 GAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGAGGCTCTCCGAAGAGAAAGAGAA 1320
QY 3794 GTGCAAACTCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3853
DB |||||
DB 1321 GTGCAAACTCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
QY 3854 CAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3913
DB |||||
DB 1381 CAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
QY 3914 GGTCTTGCACTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3973
DB |||||
DB 1441 GGTCTTGCACTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
QY 3974 GTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4033
DB |||||
DB 1501 GTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
QY 4034 AAATGAGGCGGTGGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4093
DB |||||
DB 1561 AAATGAGGCGGTGGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
QY 4094 CGAAGGAATCCATCTGGGTAA 4114

DB 1621 CGAAGGAATCCATCTCGATAA 1641
RESULT 6
US-09-949-016-4262
; Sequence 4262, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: PASTSEQ for Windows Version 4.0
; SEQ ID NO 4262
; LENGTH: 5632
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4262
Query Match 16.4%; Score 1639.4; DB 4; Length 5632;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1640; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2474 GACAAAGAGAACTAATGCTTTGCTGATTCATATTTGAATCGAGGCATTTGGAAACCTG 2533
DB |||||
DB 1 GACAAAGAGAACTAATGCTTTGCTGATTCATATTTGAATCGAGGCATTTGGAAACCTG 60
QY 2534 TATGCTCTTTGTTGTAAGAACCACTGACACCATCAGCTTCTCTAAAAGTTCGAAG 2593
DB |||||
DB 61 TATGCTCTTTGTTGTAAGAACCACTGACACCATCAGCTTCTCTAAAAGTTCGAAG 120
QY 2594 AAGTTAGAGGACTATACACTTTCTTTTGAACCTTTTATAATAATTTGCTCTGGTTTTT 2653
DB |||||
DB 121 AAGTTAGAGGACTATACACTTTCTTTGAACTTTTATAATAATTTGCTCTGGTTTTT 180
QY 2654 GAAACCCAGGCTGTAGAGGGGTGAGTGCACAGTCTTACAGTGCCTTATTCGAATC 2713
DB |||||
DB 181 GAAACCCAGGCTGTAGAGGGGTGAGTGCACAGTCTTACAGTGCCTTATTCGAATC 240
QY 2714 CAGAAATTCGCCAACCGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 2773
DB |||||
DB 241 CAGAAATTCGCCAACCGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300
QY 2774 AACTCAATCCCTCTTAATGTATCATGGAGGGCAGAGGTGATTCAGTCTCTTGGCAG 2833
DB |||||
DB 301 AACTCAATCCCTCTTAATGTATCATGGAGGGCAGAGGTGATTCAGTCTCTTGGCAG 360
QY 2834 TCCGATGGAGATGGAGGATGCCCTTGTCAATGAAGGGACCGCTGTGTTCCATTCGAGC 2893
DB |||||
DB 361 TCCGATGGAGATGGAGGATGCCCTTGTCAATGAAGGGACCGCTGTGTTCCATTCGAGC 420
QY 2894 TACACAAGAAAAATGTCTCATCAATCGAGGGGTATATGCCCTTGGATTGATGTTCTG 2953
DB |||||
DB 421 TACACAAGAAAAATGTCTCATCAATCGAGGGGTATATGCCCTTGGATTGATGTTCTG 480
QY 2954 CAGCCAGACCTTTCACACATTCAGAGACCTTAAATAACATGTCTTAATGAACACCGGCC 3013
DB |||||
DB 481 CAGCCAGACCTTTCACACATTCAGAGACCTTAAATAACATGTCTTAATGAACACCGGCC 540
QY 3014 TACCCTCTGTGAACACGAGGTTCTTCGGGTGAGCAGATATCTCAGTCCGCTTGATAA 3073
DB |||||
DB 541 TACCCTCTGTGAACACGAGGTTCTTCGGGTGAGCAGATATCTCAGTCCGCTTGATAA 600

QY 3074 AAGTCAAGTGCAGACAGACCTCCCAAGGAAAGAAATGCAAGGAAATGAAATTTAGCTG 3133
Db |||||||
QY 601 AAGTCAAGTGCAGACAGACCTCCCAAGGAAAGAAATGCAAGGAAATGAAATTTAGCTG 660
Db |||||||
QY 3134 TGAGGTATGTGGCGAGACATTTAGAGTCCGCTTTTGATGTTGAGATCCCATGAGAACACA 3193
Db |||||||
QY 661 TGAGGTATGTGGCGAGACATTTAGAGTCCGCTTTTGATGTTGAGATCCCATGAGAACACA 720
Db |||||||
QY 3194 CAAAGATCTTTTCACTTACGGGTGTAAATCATGTGCGGAAGAAATTCAGGAGCCTTGTT 3253
Db |||||||
QY 721 CAAAGATCTTTTCACTTACGGGTGTAAATCATGTGCGGAAGAAATTCAGGAGCCTTGTT 780
Db |||||||
QY 3254 TCTTAAAAATCATATGCGGACACATAATGCAAAATCGGGGCGAGAACAACTCGAGCA 3313
Db |||||||
QY 781 TCTTAAAAATCATATGCGGACACATAATGCAAAATCGGGGCGAGAACAACTCGAGCA 840
Db |||||||
QY 3314 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTGCTCCAGGTGCGACGCGCGGAGAG 3373
Db |||||||
QY 841 AGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTGCTCCAGGTGCGACGCGCGGAGAG 900
Db |||||||
QY 3374 CATCTCCTCTCTTACAAAATCTGATGTTGTTGCTTCTCTATTTTCAAATAAAGAAAG 3433
Db |||||||
QY 901 CATCTCCTCTCTTACAAAATCTGATGTTGTTGCTTCTCTATTTTCAAATAAAGAAAG 960
Db |||||||
QY 3434 TCTAAATTGAGCAGCGCAAGGTGCACACCAAAAATCTGTTTCGGTACCAGCAGCGCGCA 3493
Db |||||||
QY 961 TCTAAATTGAGCAGCGCAAGGTGCACACCAAAAATCTGTTTCGGTACCAGCAGCGCGCA 1020
Db |||||||
QY 3494 GACAGACTCTCCACAGGAGGAATGCCGTCTCGAGGGAGGACTTCCCTGCAAGTGTTC 3553
Db |||||||
QY 1021 GACAGACTCTCCACAGGAGGAATGCCGTCTCGAGGGAGGACTTCCCTGCAAGTGTTC 1080
Db |||||||
QY 3554 CTTGAGACCAAAATCTCACCTGAAACGGGGAGAGCCTGTGATGATGATCCCTCAGCT 3613
Db |||||||
QY 1081 CTTGAGACCAAAATCTCACCTGAAACGGGGAGAGCCTGTGATGATGATCCCTCAGCT 1140
Db |||||||
QY 3614 CGATCCGTTTCCACACTTCCAGCTTGGCAGCTGGCTACCAAGGAAAGTTGCCATTG 3673
Db |||||||
QY 1141 CGATCCGTTTCCACACTTCCAGCTTGGCAGCTGGCTACCAAGGAAAGTTGCCATTG 1200
Db |||||||
QY 3674 CCAAGAAGTGAAGGAATCGGGGCAAGAGGGAGCACCGCAACAACGAGTTCGAGTTCCGA 3733
Db |||||||
QY 1201 CCAAGAAGTGAAGGAATCGGGGCAAGAGGGAGCACCGCAACAACGAGTTCGAGTTCCGA 1260
Db |||||||
QY 3734 GAAGGAGCTTGAGAAAACAAATAAGGGCAGTTGTGAGGAGCTTCCGCAAGAGAGAGAA 3793
Db |||||||
QY 1261 GAAGGAGCTTGAGAAAACAAATAAGGGCAGTTGTGAGGAGCTTCCGCAAGAGAGAGAA 1320
Db |||||||
QY 3794 GTGCAACACTCCACGGGAGAGCGCCCTCCGTGGAGCGGATCCCAAGTTACCCAGTAG 3853
Db |||||||
QY 1321 GTGCAACACTCCACGGGAGAGCGCCCTCCGTGGAGCGGATCCCAAGTTACCCAGTAG 1380
Db |||||||
QY 3854 CAAGGAGAGCCCACTCACTGCTCCGAGTCCGCGCAAGAGCTTTTCAAGACCTTACCAACGCT 3913
Db |||||||
QY 1381 CAAGGAGAGCCCACTCACTGCTCCGAGTCCGCGCAAGAGCTTTTCAAGACCTTACCAACGCT 1440
Db |||||||
QY 3914 GGTCTTGCACTCAGGGTCCACAGAGGACCGAGGGCGGGCGGAGTCCGCCACCAT 3973
Db |||||||
QY 1441 GGTCTTGCACTCAGGGTCCACAGAGGACCGAGGGCGGGCGGAGTCCGCCACCAT 1500
Db |||||||
QY 3974 GTCTGTGAGCGGAGGACCGGGGAGCTTCTCTGACCTCGCGGCCCTCTGATCA 4033
Db |||||||
QY 1501 GTCTGTGAGCGGAGGACCGGGGAGCTTCTCTGACCTCGCGGCCCTCTGATCA 1560
Db |||||||
QY 4034 AAATGAGCGGTGGATCGAGGGGAGGTTGTTCTGAAGACGGATCTGAGGATGGGTTCC 4093
Db |||||||
QY 1561 AAATGAGCGGTGGATCGAGGGGAGGTTGTTCTGAAGACGGATCTGAGGATGGGTTCC 1620
Db |||||||
QY 4094 CGAAGGAATCCATCTGGGTAA 4114
Db |||||||
QY 1621 CGAAGGAATCCATCTGGGTAA 1641

RESULT 7
US-08-892-695-10
; Sequence 10, Application US/08892695A
; Patent No. 6808878
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W
; APPLICANT: Collins, Collin
; APPLICANT: Hwang, Soo In
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: GENES FROM THE 20Q13 AMPLICON AND THEIR USES
; FILE REFERENCE: 2500.124US3
; CURRENT APPLICATION NUMBER: US/08/892,695A
; CURRENT FILING DATE: 1997-07-15
; EARLIER APPLICATION NUMBER: 08/785,532
; EARLIER FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/731,499
; EARLIER FILING DATE: 1996-10-16
; EARLIER APPLICATION NUMBER: 08/680,395
; EARLIER FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 3186
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: ZABCI Open
; OTHER INFORMATION: Reading Frame
US-08-892-695-10

Query Match 15.6%; Score 1559.6; DB 4; Length 3186;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1565; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 8286 AGGTCAAAAACCATCAAAATGGAATTTTGTGAATATGTCAGGCCAGAACATCTCT 8345
Db |||||||
QY 1482 AGGTCAAAAACCATCAAAATGGAATTTTGTGAATATGTCAGGCCAGAACATCTCT 1541
Db |||||||
QY 8346 GAGGTATCACTTGGAGAGACATCAAGGAAAAAACAACCGATGTTGCTGAGAGTCAA 8405
Db |||||||
QY 1542 GAGGTATCACTTGGAGAGACATCAAGGAAAAAACAACCGATGTTGCTGAGAGTCAA 1601
Db |||||||
QY 8406 GACCATGGTAAATCAGGACACTGAAGATGCACTATTAAACCGCTGACAGTCCCAAC 8465
Db |||||||
QY 1602 GAACCATGGTAAATCAGGACACTGAAGATGCACTATTAAACCGCTGACAGTCCCAAC 1661
Db |||||||
QY 8466 CAAAAATTTGAAAAGATTTTGTGATGGTGCCAAAGATGTTACAGGAGTCCACCTGCMAA 8525
Db |||||||
QY 1662 CAAAAATTTGAAAAGATTTTGTGATGGTGCCAAAGATGTTACAGGAGTCCACCTGCMAA 1721
Db |||||||
QY 8526 GCAGCTTAAGGAGATGCTTCTGTTTTTTCAGAAATGTTCTGGGAGCGCTGCTCTCAC 8585
Db |||||||
QY 1722 GCAGCTTAAGGAGATGCTTCTGTTTTTTCAGAAATGTTCTGGGAGCGCTGCTCTCAC 1781
Db |||||||
QY 8586 AGCACAAAAGATCTCAGGATTTCCATAAAAATCAGCTGATGATGCTGATATAAGT 8645
Db |||||||
QY 1782 AGCACAAAAGATCTCAGGATTTCCATAAAAATCAGCTGATGATGCTGATATAAGT 1841
Db |||||||
QY 8646 GAATAAAAACCTTACCCCTGCTTACTCGGACCTGTTTAAAAAAGAGATCAGCAGTGAAC 8705
Db |||||||
QY 1842 GAATAAAAACCTTACCCCTGCTTACTCGGACCTGTTTAAAAAAGAGATCAGCAGTGAAC 1901
Db |||||||
QY 8706 TCAGGCAATAAACCCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCGATGGCAG 8765
Db |||||||
QY 1902 TCAGGCAATAAACCCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCGATGGCAG 1961
Db |||||||
QY 8766 TACCAACCACTTGAAGTTAGCCCAAGAGAGCAAAACCGAGAGCCGAGCTGACCTG 8825
Db |||||||
QY 1962 TACCAACCACTTGAAGTTAGCCCAAGAGAGCAAAACCGAGAGCCGAGCTGACCTG 2021
Db |||||||
QY 8826 CAGATACAGGCCCAAGTGTGGATTGTACGAAAAACCTTTAAATTTTATCCGTGGGGCTCT 8885

Db 2022 CAGATACAGCCCAAGTGTGATTGTACGAAACCTTTAAATTTATCCGTGGGGCTCT 2081
QY 8886 TCACAAATGCGCGCAATTTCTTTGAGTAAAGTTTGGATTTCAAGTATACCTGTCCATT 8945
Db 2082 TCACAAATGCGCGCAATTTCTTTGAGTAAAGTTTGGATTTCAAGTATACCTGTCCATT 2141
QY 8946 TTGTACTTCAAGACATTTTATCCAGAAATTTTAAATGATCCACAGAGACTGGAGCATAA 9005
Db 2142 TTGTACTTCAAGACATTTTATCCAGAAATTTTAAATGATCCACAGAGACTGGAGCATAA 2201
QY 9006 ATACAAATCTGAGCTTCAATAAACTGTGCAAAAGTCTTCTTTAGAGTGCAGTAC 9065
Db 2202 ATACAAATCTGAGCTTCAATAAACTGTGCAAAAGTCTTCTTTAGAGTGCAGTAC 2261
QY 9066 CGGATGCGCCCGCAGCGTGTGCGAAAGATGTGCTCCCTCCCTAGTTTCTGTAAACC 9125
Db 2262 CGGATGCGCCCGCAGCGTGTGCGAAAGATGTGCTCCCTCCCTAGTTTCTGTAAACC 2321
QY 9126 CAAGCCCAAGTCTGCTTTCCCGCGCAGTCCAAATCCCTGCGATCTGCGAAGGGGAAGCA 9185
Db 2322 CAAGCCCAAGTCTGCTTTCCCGCGCAGTCCAAATCCCTGCGATCTGCGAAGGGGAAGCA 2381
QY 9186 GAGCCCTCTGGCGCAGGCGAGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGC 9245
Db 2382 GAGCCCTCTGGCGCAGGCGAGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGC 2441
QY 9246 CCCAAGTAACTGCAAGTCCACAGACACAGCAGAAATGTGGGGTCCAGGGGCCGCCAC 9305
Db 2442 CCCAAGTAACTGCAAGTCCACAGACACAGCAGAAATGTGGGGTCCAGGGGCCGCCAC 2501
QY 9306 CAGGCAACGAGCAATCTGAGATGTTTCTTAAACAGTGTTCCTGCGACCGGATAGAC 9365
Db 2502 CAGGCAACGAGCAATCTGAGATGTTTCTTAAACAGTGTTCCTGCGACCGGATAGAC 2561
QY 9366 AAAAGACCCGAGACAAAATTTGAACTCTTCAGTAGTCTCTTCAGCCACCTCGG 9425
Db 2562 AAAAGACCCGAGACAAAATTTGAACTCTTCAGTAGTCTCTTCAGCCACCTCGG 2621
QY 9426 CAGCAGTAACTCAATGTTTCCATCGACTACCCCGCAGAGACGACGCGGTGGGCACC 9485
Db 2622 CAGCAGTAACTCAATGTTTCCATCGACTACCCCGCAGAGACGACGCGGTGGGCACC 2681
QY 9486 TCCGGAAGAGACTATTTCTGTAATCGAGTGCAGCAATCTGCGACGAGAAATTTGGTGA 9545
Db 2682 TCCGGAAGAGACTATTTCTGTAATCGAGTGCAGCAATCTGCGACGAGAAATTTGGTGA 2741
QY 9546 GCGCCTTCCAAAAGACTGAAGTCCAGCGTGTGCTTCCATCGACTACCCCGCAGAGACGACGCGGTGGGCACC 9605
Db 2742 GCGCCTTCCAAAAGACTGAAGTCCAGCGTGTGCTTCCATCGACTACCCCGCAGAGACGACGCGGTGGGCACC 2801
QY 9606 CAATTACAGAGGCTATGACCTTCCCAAGTACCATATGCTCAGAGGATCACATCACT 9665
Db 2802 CAATTACAGAGGCTATGACCTTCCCAAGTACCATATGCTCAGAGGATCACATCACT 2861
QY 9666 GTTACCGCAGGACTGTGTATCGCTGCGAGGCGCTGCTCCCAACCAAGTTCCTGAG 9725
Db 2862 GTTACCGCAGGACTGTGTATCGCTGCGAGGCGCTGCTCCCAACCAAGTTCCTGAG 2921
QY 9726 CTCAGCGAGGCTGATTTCCAAATGCTGACTTTCAGAGCCCTATGCTGGTCCGG 9785
Db 2922 CTCAGCGAGGCTGATTTCCAAATGCTGACTTTCAGAGCCCTATGCTGGTCCGG 2981
QY 9786 GGCACCTTACACTTGTGCTGCTGCTGAGTCCAGCATCCAGCTCGACCTTGAAGGATAT 9845
Db 2982 GGCACCTTACACTTGTGCTGCTGCTGAGTCCAGCATCCAGCTCGACCTTGAAGGATAT 3041
QY 9846 TGCATGAGGGGGT 9859
Db 3042 TGTGATGTCAGT 3055

US-08-680-395-3
; Sequence 3, Application US/08680395
; Patent No. 5892010
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W.
; APPLICANT: Collins, Colin
; APPLICANT: Hwang, Soo-in
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowbel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: Genes from the 20q13 Amplicon and Their
; TITLE OF INVENTION: Uses
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/680,395
; FILING DATE: 15-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-068900US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1507 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..1507
; OTHER INFORMATION: /note= "cDNA clone cc49 of 6-7kb
; OTHER INFORMATION: transcript with homology to C2H2 zinc
; OTHER INFORMATION: finger genes"
US-08-680-395-3

Query Match 13.9%; Score 1387.2; DB 2; Length 1507;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1470; Conservative 3; Mismatches 32; Indels 6; Gaps 6;
QY 2424 CAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAAGAGA 2483
Db 2 CAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAAGAGC 61
QY 2484 ACTAATGCTTGTGCTGATTCATATTTGAATCGAGCATTTGGAAACCTGTATGCTTGT 2543
Db 62 ACTAATGCTTGTGCTGATTCATATTTGAATCGAGGCATTTGGAAACCTGTATGCTTGT 121
QY 2544 TTGTGGAAAGACCAAGTGCACCATCAGCTTCTTAAAGTTTCGAGAGAGTTAGAGG 2603
Db 122 TTGTGGAAAGACCAAGTGCACCATCAGCTTCTTAAAGTTTCGAGAGAGTTAGAGG 181
QY 2604 ACTATACACTTCTTTTGAACCTTTTATAATAATATTTGCTCTGCTGTTTGGAAACCCAGG 2663
Db 182 ACTATACACTTCTTTTGAACCTTTTATAATAATATTTGCTCTGCTGTTTGGAAACCCAGG 240
QY 2664 GCTGTTAGAGGGGTGAGTGACAAAGTCTTTACAGTGGCTTTATTCCAACTCCAGAAATTGC 2723

Db 241 ACTGTTAGA-GGGTGAGTGACAGGCTCTTAC-AGTGGCCTTAAATCCAACTCCAGAAATTGC 298
Qy 2724 CCAACGGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATCCCAACTCAATCC 2783
Db 299 CCAACGGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 358
Qy 2784 CTCCTTAATGTACATGGATGGGCCAAGAGTATGTCAGCTCTCTTGGCAGTCCGATGAG 2843
Db 359 CTCCTTAATGTACATGGAT-GGCCAAGAGTATGTCAGCTCTCTTGGCAGTCCGATGAG 417
Qy 2844 ATGAGGAGTCCCTCTCAATGAAGGGACCGCTGTGTTGTTCCATTCCGAGCTACACAAGAA 2903
Db 418 ATGGA-GATGCTCTGTCATGAAGAGGCCNCCTGTTGTCAATTCGAGCTACACAAGAA 476
Qy 2904 AAAAATGTC-ATCCAAATCGAGGGGTATATGCCCCTTGGATTGCAATGTTCTGCAGCCAGAC 2962
Db 477 AAAAATGTC-ATCCAAATCGAGGGGAATATGCCCCTTGGATTGCAATGTTCTGCAGCCAGAC 536
Qy 2963 CTTTCACATTCAGAAGACCTTAAATAACATGTCTTAATGCAACACCGCCCTACCCCTCTG 3022
Db 537 CTTTCACATTCAGAAGACCTTAAATAACATGTCTTAATGCAACACCGCCCTACCCCTCTG 596
Qy 3023 TGAACAGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCGCTTGATGAAGTCAAGT 3082
Db 597 TGAACAGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCGCTTGATGAAGTCAAGT 656
Qy 3083 GCGAACAGAACCTCCCAAGGAAAGAAATGCAAGGAAATGAATTTAGCTGTGAGGTATG 3142
Db 657 GCGAACAGAACCTCCCAAGGAAAGAAATGCAAGGAAATGAATTTAGCTGTGAGGTATG 716
Qy 3143 TGGGAGACATTTAGAGTCGCTTTTGAATGTTGAGATCCACATGAGAACACACAAAGATTC 3202
Db 717 TGGGAGACATTTAGAGTCGCTTTTGAATGTTGAGATCCACATGAGAACACACAAAGATTC 776
Qy 3203 TTTCTACTACGGGTGTAACTGTGCGAAGAGATTCAAGGAGCCTTGTTCTTTAAAAA 3262
Db 777 TTTCTACTACGGGTGTAACTGTGCGAAGAGATTCAAGGAGCCTTGTTCTTTAAAAA 836
Qy 3263 TCACATCGGACACATAATGGCAAAATCGGGGCCAGAGCAAACTGCAGCAAGGCTTGA 3322
Db 837 TCACATCGGACACATAATGGCAAAATCGGGGCCAGAGCAAACTGCAGCAAGGCTTGA 896
Qy 3323 GAGTAGTCAGCAACGATCAACAGGTCGTCCAGGTGCAACGGCGGAGAGCATCTCTC 3382
Db 897 GAGTAGTCAGCAACGATCAACAGGTCGTCCAGGTGCAACGGCGGAGAGCATCTCTC 956
Qy 3383 TCCTTACAAAATCTGCATGGTTGTGGCTTCTCTATTCCAAAATAAGAAAGTCTAATTGA 3442
Db 957 TCCTTACAAAATCTGCATGGTTGTGGCTTCTCTATTCCAAATAAGAAAGTCTAATTGA 1016
Qy 3443 GCACCGCAAGGTGCACACAAAAAACTGCTTTTCGGTACCAGCGCGGAGAGCATCTC 3502
Db 1017 GCACCGCAAGGTGCACACAAAAAACTGCTTTTCGGTACCAGCGCGGAGAGCATCTC 1076
Qy 3503 TCACAGAGGAAATGCGCTCTCGAGGAGACTTCTCGAGTGTTCACCTTCAGACC 3562
Db 1077 TCACAGAGGAAATGCGCTCTCGAGGAGACTTCTCGAGTGTTCACCTTCAGACC 1136
Qy 3563 AAAATCTCACCTCGAAAACGGGAAAGAGCTGTCCAGATGCAATCCCTCAGCTCGATCCGTT 3622
Db 1137 AAAATCTCACCTCGAAAACGGGAAAGAGCTGTCCAGATGCAATCCCTCAGCTCGATCCGTT 1196
Qy 3623 CACCACTTCAGGCTTGGCAGCTACCAAGGAAAGTTCATTTGCAAGTTCAGAAAGT 3682
Db 1197 CACCACTTCAGGCTTGGCAGCTACCAAGGAAAGTTCATTTGCAAGTTCAGAAAGT 1256
Qy 3683 GAGGAAATCGGGCAGAGGAGGACCGACACACGATTCGAGTTCGAGAGGAGCT 3742
Db 1257 GAGGAAATCGGGCAGAGGAGGACCGACACACGATTCGAGTTCGAGAGGAGCT 1316
Qy 3743 TGGAGAAACAAATAAGGGCAGTTGTGCGAGGCTCTCGCAAGAGAAAGAGTGCACAA 3802
Db 1317 TGGAGAAACAAATAAGGGCAGTTGTGCGAGGCTCTCGCAAGAGAAAGAGTGCACAA 1376

Qy 3803 CTCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAGCAAGAGAA 3862
Db 1377 CTCCACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAGCAAGAGAA 1436
Qy 3863 GCCCACTCAGTCTCCGAGTGGCGGCAAGCTTTTCAGAACCTTACCAACAGCTGGTCTTGA 3922
Db 1437 GCCCACTCAGTCTCCGAGTGGCGGCAAGCTTTTCAGAACCTTACCAACAGCTGGTCTTGA 1496
Qy 3923 CTCCAGGCTCC 3933
Db 1497 CTCCAGGCTCC 1507

RESULT 9

US-08-892-695-3
; Sequence 3, Application US/08892695A
; Patent No. 6808878
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W
; APPLICANT: Collins, Collin
; APPLICANT: Hwang, Soo In
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: GENES FROM THE 20Q13 AMPLICON AND THEIR USES
; FILE REFERENCE: 2500.124US3
; CURRENT APPLICATION NUMBER: US/08/892,695A
; CURRENT FILING DATE: 1997-07-15
; EARLIER APPLICATION NUMBER: 08/785,532
; EARLIER FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/731,499
; EARLIER FILING DATE: 1996-10-16
; EARLIER APPLICATION NUMBER: 08/680,395
; EARLIER FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1507
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cc49
; NAME/KEY: modified_base
; LOCATION: (447)
; OTHER INFORMATION: N is A, G, C, T, or U
US-08-892-695-3

Query Match 13.9%; Score 1387.2; DB 4; Length 1507;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1470; Conservative 3; Mismatches 32; Indels 6; Gaps 6;
Qy 2424 CAGGTTGCTGGATTGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAAGAGA 2483
Db 2 CAGGTTGCTGGATTGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAAGAGC 61
Qy 2484 ACTAATGCTTTTGTGCTGATTCAATTTGAATCGAGGCATTTGGAAACCTGTATGCTTGT 2543
Db 62 ACTAATGCTTTTGTGCTGATTCAATTTGAATCGAGGCATTTGGAAACCTGTATGCTTGT 121
Qy 2544 TTGTGGAAGAACCAAGTGCACCATCAGCTTCTTAAAGTTTCGAGAGTTAGAGG 2603
Db 122 TTGTGGAAGAACCAAGTGCACCATCAGCTTCTTAAAGTTTCGAGAGTTAGAGG 181
Qy 2604 ACTATACACTTTCTTTTGAACCTTTTATATAATATTTGCTCTGCTTTTGGAAACCCAGG 2663
Db 182 ACTATACACTTTCTTTTGAACCTTTTATATAATATTTGCTCTGCTTTTGGAAACCCAGG 240
Qy 2664 GCTGTTAGAGGGGTGAGTGACAAAGTCTTTACAGTGGCTTTATTCACACTCCAGAAATTGC 2723
Db 241 ACTGTTAGA-GGGTGAGTGACAGGCTTTAC-AGTGGCCTTAAATCCAACTCCAGAAATTGC 298

QY 2724 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAACATGCGCAACTCAATCC 2783
DB CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAACATGCGCAACTCAATCC 358
QY 2784 CTCTTAATGTACATGATGGCCAGAGTGAATTTGGCAGCTCTCTTGGCAGTCCGATGGAG 2843
DB CTCTTAATGTACATGATGGCCAGAGTGAATTTGGCAGCTCTCTTGGCAGTCCGATGGAG 417
QY 2844 ATGGAGGATGCTTGTCAATGAAAGGACCGCTGTGTTCATTCGAGCTACACAAGAA 2903
DB ATGGA-GATGCTTGTCAATGAAAGGACCGCTGTGTTCATTCGAGCTACACAAGAA 476
QY 2904 AAAAATGTC-ATCCAAATCGAGGGTATATGCGCTTTGGATGTGATGCTTCGAGCCAGAC 2962
DB AAAAATGTC-ATCCAAATCGAGGGTATATGCGCTTTGGATGTGATGCTTCGAGCCAGAC 536
QY 2963 CTTTACACATTCAGAGACCTTAATAACATGCTTAATGCAACACCGGCTTACCCTCTG 3022
DB CTTTACACATTCAGAGACCTTAATAACATGCTTAATGCAACACCGGCTTACCCTCTG 596
QY 3023 TGAACGAGCAGTCTTTCGGGTGAAGCAGAGTATCTCAGTCCGCTTGTATATAAAGTCAAGT 3082
DB TGAACGAGCAGTCTTTCGGGTGAAGCAGAGTATCTCAGTCCGCTTGTATATAAAGTCAAGT 556
QY 3083 GCGAACAGAACCTCCCAAGGAAAGAAATGCAAGGAAATGAATTTAGCTGTGAGTATG 3142
DB GCGAACAGAACCTCCCAAGGAAAGAAATGCAAGGAAATGAATTTAGCTGTGAGTATG 716
QY 3143 TGGGCAGACATTTAGAGTGCCTTTTGAATGTTGAGATCCATGAGAACACACAAGATTC 3202
DB TGGGCAGACATTTAGAGTGCCTTTTGAATGTTGAGATCCATGAGAACACACAAGATTC 776
QY 3203 TTTTCACTTACGGGTGTAACATGTGCGAAGAGATTTCAAGGACCTTGTGTTCTTAAAAA 3262
DB TTTTCACTTACGGGTGTAACATGTGCGAAGAGATTTCAAGGACCTTGTGTTCTTAAAAA 836
QY 3263 TCACATGCGGACACATAATGCGAAATCGGGGGCAGAAGCAAACTCAGCAAGGCTTGA 3322
DB TCACATGCGGACACATAATGCGAAATCGGGGGCAGAAGCAAACTCAGCAAGGCTTGA 896
QY 3323 GAGTAGTCAGACACATCAACGAGTGTCTCAGGTGCGACGGCGCAGAGCATCTCCTC 3382
DB GAGTAGTCAGACACATCAACGAGTGTCTCAGGTGCGACGGCGCAGAGCATCTCCTC 956
QY 3383 TCCTTCAAAAATCTGATGTTTGGCTTCTTATTTCCAAATPAAAGAAAGTCTAATTCGA 3442
DB TCCTTCAAAAATCTGATGTTTGGCTTCTTATTTCCAAATPAAAGAAAGTCTAATTCGA 1016
QY 3443 GCACCGCAAGGTGCACACCAAAAAAACTGCTTTCCGGTACCAGCGCGCAGACAGCTC 3502
DB GCACCGCAAGGTGCACACCAAAAAAACTGCTTTCCGGTACCAGCGCGCAGACAGCTC 1076
QY 3503 TCCACAGAGGAGATGCGCTCTCGAGGAGGAGCTTCTCGAGTGTGTTCAACTTGAGACC 3562
DB TCCACAGAGGAGATGCGCTCTCGAGGAGGAGCTTCTCGAGTGTGTTCAACTTGAGACC 1076
QY 3563 AAAATCTCACTCCCTGAAACGGGGAAGAGCTGTGAGATGATCCCTCAGCTCGATCCGTT 3622
DB AAAATCTCACTCCCTGAAACGGGGAAGAGCTGTGAGATGATCCCTCAGCTCGATCCGTT 1196
QY 3623 CACCACTTTCCAGGCTTGGCAGTGGCTTACCAAGGAAAGTTGCCATTTGCGCAAGAGT 3682
DB CACCACTTTCCAGGCTTGGCAGTGGCTTACCAAGGAAAGTTGCCATTTGCGCAAGAGT 1256
QY 3683 GAAGGATCGGGCAAGAGGAGGACCCGACAAAGAGATTCGAGTTCGAGAGAGGCT 3742
DB GAAGGATTCGGGCAAGAGGAGGACCCGACAAAGAGATTCGAGTTCGAGAGAGGCT 1316
QY 3743 TCGAGAAACAAATAAGGGCAGTTGTGAGGCTCTCGCAAGAGAGAGAGTGCAGACA 3802
DB TCGAGAAACAAATAAGAGACATTTGTGAGGCTCTCGCAAGAGAGAGAGTGCAGACA 1376
QY 3803 CTCCACGCGGAAGCGCCCTCGGTGAGCGCGGATCCCAAGTTTACCAGTAGCAAGAGAA 3862

DB 1377 CTCCACGCGGAAGCGCCCTCGGTGACGCGGATCCCAAGTTACCCAGTAGCAAGGAGAA 1436
QY 3863 GCCCACTCACTGCTCCGAGTGGCGGCAAGCTTTAGAACCTTACACAGCTGCTTTCGA 3922
DB 1437 GCCCACTCACTGCTCCGAGTGGCGGCAAGCTTTAGAACCTTACACAGCTGCTTTCGA 1496
QY 3923 CTCACGGGTCC 3933
DB 1497 CTCACGGGTCC 1507
RESULT 10
US-09-949-016-34699/c
; Sequence 34699, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34699
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-34699
Query Match 6.0%; Score 600.6; DB 4; Length 601;
Best Local Similarity 99.8%; Pred. No. 8.1e-141;
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2843 GATGAGGATGCTTGTCAATGAAAGGGACCGCTGTGTTCATTCGAGCTACACAAGA 2902
DB 601 GATGAGGATGCTTGTCAATGAAAGGGACCGCTGTGTTCATTCGAGCTACACAAGA 542
QY 2903 AAAAATGTCAATCGAGGGTATATGCGCTTGGATTCGATGCTTCGAGCCAGAC 2962
DB 541 AAAAATGTCAATCGAGGGTATATGCGCTTGGATTCGATGCTTCGAGCCAGAC 482
QY 2963 CTTTACACATTCAGAGACCTTTAATAAACATGCTTTAATGCAACACCGGCTTACCCTCTG 3022
DB 481 CTTTACACATTCAGAGACCTTTAATAAACATGCTTTAATGCAACACCGGCTTACCCTCTG 422
QY 3023 TGAACGAGCAGTCTTTCGGGTGAAGCAGAGTATCTCAGTCCGCTTGTATATAAAGTCAAGT 3082
DB 421 TGAACGAGCAGTCTTTCGGGTGAAGCAGAGTATCTCAGTCCGCTTGTATATAAAGTCAAGT 362
QY 3083 GCGAACAGAACCTCCCAAGGAAAGAAATGCGAAGGAAATGAATTTAGCTGTGAGGTATG 3142
DB 361 GCGAACAGAACCTCCCAAGGAAAGAAATGCGAAGGAAATGAATTTAGCTGTGAGGTATG 302
QY 3143 TGGGCAGACATTTAGAGTGCCTTTTGAATGTTGAGATCCACATGAGAACACACAAGATTC 3202
DB 301 YGGGCAGACATTTAGAGTGCCTTTTGAATGTTGAGATCCACATGAGAACACACAAGATTC 242
QY 3203 TTTTCACTTACGGGTGTAACATGTGCGGGAAGAGATTTCAAGGAGCCTTGTGTTCTTAAAAA 3262
DB 241 TTTTCACTTACGGGTGTAACATGTGCGGGAAGAGATTTCAAGGAGCCTTGTGTTCTTAAAAA 182
QY 3263 TCACATGCGGACACATAATGCGAAATCGGGGGCAGAAGCAAACTCAGCAAGGCTTGA 3322
DB 181 TCACATGCGGACACATAATGCGAAATCGGGGGCAGAAGCAAACTCAGCAAGGCTTGA 122

Db 181 CCCAGCTTGGCGTTGAGTTCTGGTTCTACCACTGGCTGTTTGTGACCAATATGAGT 122
QY 5730 TGCTTAACCTTTCTTTGCTACTATTTCCCTGTTTGCATAATGCTTCAATGACCCCTGTCT 5789
Db 121 TGCTTAACCTTTCTTTGCTACTATTTCCCTGTTTGCATAATGCTTCAATGACCCCTGTCT 62
QY 5790 TCCACCTCCCAAGGACAAATTTCAACAGCTTATTTGTAAAGATCAAGTCTTTAAAAA 5849
Db 61 TCCACCTCCCAAGGACAAATTTCAACAGCTTATTTGTAAAGATCAAGTCTTTAAAAA 2
QY 5850 A 5850
Db 1 A 1

RESULT 13
US-09-949-016-34702/c
; Sequence 34702, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34702
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-34702

Query Match 6.0%; Score 600.6; DB 4; Length 601;
Best Local Similarity 99.8%; Pred. No. 8.1e-141;
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 7580 AGTACTCTCCACCAACCGGCTAATTTCTGTATTTTGTAGAGATCGGGTTTCCACCA 7639
Db 601 AGTACTCTCCACCAACCGGCTAATTTCTGTATTTTGTAGAGATCGGGTTTCCACCA 542
QY 7640 TCGTGGCCAGGCTGTTTCGAACTCTCTGACCTCAAGTGATCCGCTCACCTTGGCCCTCCA 7699
Db 541 TCGTGGCCAGGCTGTTTCGAACTCTCTGACCTCAAGTGATCCGCTCACCTTGGCCCTCCA 482
QY 7700 TAGTGTGGCTCCCATAGTGTGGGATTAAGGCTGAGCCACCGCCCGGACAAAGT 7759
Db 481 TAGTGTGGCTCCCATAGTGTGGGATTAAGGCTGAGCCACCGCCCGGACAAAGT 422
QY 7760 TCATTTGTTTGTATGATGCTATGCTGCTGCTATCTTATTAAGAGCTACAGTAT 7819
Db 421 TCATTTGTTTGTATGATGCTATGCTGCTGCTATCTTATTAAGAGCTACAGTAT 362
QY 7820 TTTAAATGCTGCATCTTATGCTTTTATGATTGAGAAATGAAATCTATTAGTAG 7879
Db 361 TTTAAATGCTGCATCTTATGCTTTTATGATTGAGAAATGAAATCTATTAGTAG 302
QY 7880 TCTTGAGATGTGAAGAGAGCTATGACATCATGATGAGAGCTGCGTAGATTTGAAT 7939
Db 301 YCTTGAGATGTGAAGAGAGCTATGACATCATGATGAGAGCTGCGTAGATTTGAAT 242
QY 7940 TTTCACTCTTCCACTTACTATCTGTGACCCCTGGGCAAGTTATTTAACTTTTGTGCT 7999
Db 241 TTTCACTCTTCCACTTACTATCTGTGACCCCTGGGCAAGTTATTTAACTTTTGTGCT 182
QY 8000 TTTAGTTTCTTGTGTTAAAGTAGAATAATACATATTTCCCTAGGCGTGTAGGAAGA 8059

Db 181 TTTAGTTTCTTGTGTTAAAGTAGAATAATACATATTTCCCTAGGCGTGTAGGAAGA 122
QY 8060 TTTAAATAAGTTAGAAGTGTGCTGTTAAATTTTCTTATGAAGATAGGCATTTCAATTTTC 8119
Db 121 TTTAAATAAGTTAGAAGTGTGCTGTTAAATTTTCTTATGAAGATAGGCATTTCAATTTTC 62
QY 8120 AAATATTTCATTACAGTAAGGATGATAAAGAACTGATGAGAAATCCTATGTGATAGTAGAT 8179
Db 61 AAATATTTCATTACAGTAAGGATGATAAAGAACTGATGAGAAATCCTATGTGATAGTAGAT 2
QY 8180 C 8180
Db 1 C 1

RESULT 14
US-09-949-016-151879/c
; Sequence 151879, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151879
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-151879

Query Match 6.0%; Score 600.6; DB 4; Length 601;
Best Local Similarity 99.8%; Pred. No. 8.1e-141;
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2843 GATGAGGATGCTTGTCTAATCAAGAGGACCCGCTGTTTCCATTCGGAGCTACACAAGA 2902
Db 601 GATGAGGATGCTTGTCTAATCAAGAGGACCCGCTGTTTCCATTCGGAGCTACACAAGA 542
QY 2903 AAAAAATGTCATCCAAATCGAGGGTATATGCCCTTGGATTGTCATGTTCTGCAAGCCAGAC 2962
Db 541 AAAAAATGTCATCCAAATCGAGGGTATATGCCCTTGGATTGTCATGTTCTGCAAGCCAGAC 482
QY 2963 CTTTCACATTCAGAAGACCTTAAATAACATGCTTTAAATGCAACACCGGCTTACCCTCTG 3022
Db 481 CTTTCACATTCAGAAGACCTTAAATAACATGCTTTAAATGCAACACCGGCTTACCCTCTG 422
QY 3023 TGAACACGAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCGCTTGATGATAAAGTCAAGT 3082
Db 421 TGAACACGAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCGCTTGATGATAAAGTCAAGT 362
QY 3083 GCGAACAGAACCTCCCAAGGAAAAGAAATGCAAGGAAAATGAATTTAGCTGTGAGGTATG 3142
Db 361 GCGAACAGAACCTCCCAAGGAAAAGAAATGCAAGGAAAATGAATTTAGCTGTGAGGTATG 302
QY 3143 TGGCGAGACATTTAGAGTCGCTTTTGTGTTGAGATCCACATGAGAACACACAAAGATTC 3202
Db 301 YGGCGAGACATTTAGAGTCGCTTTTGTGTTGAGATCCACATGAGAACACACAAAGATTC 242
QY 3203 TTTCTACTTACGGGTGTACATGTGCGGAAGAAAGATTTCAAGGAGCCCTTGTGTTCTTAAAAA 3262
Db 241 TTTCTACTTACGGGTGTACATGTGCGGAAGAAAGATTTCAAGGAGCCCTTGTGTTCTTAAAAA 182

3263	TCACATCGGACACATAATGGCAAAATCGGGGGCCAGAGCAAACTGCAGCAAGGCTTGGA	3322
Db	TCACATCGGACACATAATGGCAAAATCGGGGGCCAGAGCAAACTGCAGCAAGGCTTGGA	122
3323	GAGTAGTCCAGCAACGATCAACGAGGTCTGTCAGGTGCACGCGCGCGAGGAGCATCTCCCTC	3382
Db	GAGTAGTCCAGCAACGATCAACGAGGTCTGTCAGGTGCACGCGCGCGAGGAGCATCTCCCTC	62
3383	TCCTTACAAAATCTGCATGGTTGTGGCTTCTCTATTTCCAAATAAAGAAAGTCTAATTGA	3442
Db	TCCTTACAAAATCTGCATGGTTGTGGCTTCTCTATTTCCAAATAAAGAAAGTCTAATTGA	2
3443	G 3443	
Db	1 G 1	

```

RESULT 15
US-09-949-016-151880/c
; Sequence 151880, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0601307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151880
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-151880

```

Query Match	6.0%;	Score	600.6;	DB	4;	Length	601;
Best Local Similarity	99.8%;	Pred.	No. 8.1e-141;				
Matches	600;	Conservative	1;	Mismatches	0;	Indels	0;
Qy	3470	TGCTTTTCGGTACCAAGCGCGCAGACAGACTCTCCACAAGGAGGAATCCGCTCCTCGAG	3529				
Db	601	TGCTTTTCGGTACCAAGCGCGCAGACAGACTCTCCACAAGGAGGAATCCGCTCCTCGAG	542				
Qy	3530	GGAGGACTTCTCTGCAGTTGTTCAACTTTGAGACCAAAATCTCACCCCTGAAAACGGGGGAAGAA	3589				
Db	541	GGAGGACTTCTCTGCAGTTGTTCAACTTTGAGACCAAAATCTCACCCCTGAAAACGGGGGAAGAA	482				
Qy	3590	GCCTGTTCAGATCGATCCCTCAGCTCGATCCGTTTCAACCACCTTCCAGAGCTTTGGCAGCTGGC	3649				
Db	481	GCCTGTTCAGATCGATCCCTCAGCTCGATCCGTTTCAACCACCTTCCAGAGCTTTGGCAGCTGGC	422				
Qy	3650	TACCAAGGAAAAGTTGTCATTTGCCAAGAAGTCAAGGAATTCGGGGCAAGAAAGGGAGCAC	3709				
Db	421	TACCAAGGAAAAGTTGCCATTTGCCAAGAAGTCAAGGAATTCGGGGCAAGAAAGGGAGCAC	362				
Qy	3710	CGACAAACGACGATTCGAGTTCCGAGAGAGGAGCTTTGGAGAAACAAATTAAGGGCAGTTGTGC	3769				
Db	361	CGACAAACGACGATTCGAGTTCCGAGAGAGGAGCTTTGGAGAAACAAATTAAGGGCAGTTGTGC	302				
Qy	3770	AGGCTCTTCGCAACGAGAAAGAGAGTGCAAACACTCCCAACGGCGAAGCGCCCTCCGTTGGA	3829				
Db	301	WGGGCTCTTCGCAACGAGAAAGAGAGTGCAAACACTCCCAACGGCGAAGCGCCCTCCGTTGGA	242				
Qy	3830	CGCGGATCCCCAAGTTACCCAGTAGCAAGGAGGAAGCCCACTCACTGCTCCGAGTGCGGCAA	3889				
Db	241	CGCGGATCCCCAAGTTACCCAGTAGCAAGGAGGAAGCCCACTCACTGCTCCGAGTGCGGCAA	182				

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 09:38:42 ; Search time 3126 Seconds
(without alignments)
18937.112 Million cell updates/sec

Title: US-08-731-499-9_COPY_1_10000
Perfect score: 10000
Sequence: 1 CCATCATATTTCTTATTTT.....ATTCTTAGATACGCCAGTGG 10000

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	9980.6	99.8	36022	11	ACN44986	Acn44986 Human gen
2	9551.6	96.5	10282	2	AAV09023	AAV09023 Homo sapi
3	6382.2	63.8	14906	4	AAK81093	AAK81093 Human imm
4	1639.4	16.4	5632	8	ACC50334	Acc50334 Breast ca
5	1639.4	16.4	5632	10	ADL14635	Add14635 Human src
6	1639.4	16.4	5632	13	ADR66761	Adr66761 Human pro
7	1639.4	16.4	5632	13	ADR65858	Adr65858 Human pro
8	1639.4	16.4	5632	12	ADP07267	Adp07267 Human ZNF
9	1633.4	16.3	1797	4	AAK61822	AAK61822 Human imm
10	1559.6	15.6	5252	11	ACN44987	Acn44987 Human mRN
11	1554.8	15.5	3183	2	AAV09024	AAV09024 Homo sapi
12	1387.2	13.9	1507	2	AAV04698	AAV04698 Homo sapi
13	1306.8	10.1	124987	12	ADP03055	Adp03055 Human hou
14	1006.8	10.1	124990	13	ADS88553	Ads88553 Human hou
15	820	8.2	26345	11	ACN44984	Acn44984 Mouse gen
16	726.4	7.3	3016	11	ACN44985	Acn44985 Mouse mRN
17	566	5.7	566	4	AAK81095	AAK81095 Human imm
18	530	5.3	530	12	ACH76516	Ach76516 Human gen
19	335.8	3.4	469	2	AAV88042	AAV88042 EST clone
20	256.8	2.6	501	12	ACH67097	Ach67097 Human gen

C	21	251	2.5	251	12	ACH90216	Ach90216 Human gen
C	22	233.8	2.3	37314	4	AAK71358	AAK71358 Human imm
C	23	233.8	2.3	172570	6	ABQ88207	Abq88207 Human ost
C	24	227.2	2.3	104245	12	ADG86768	Adg86768 Human clo
C	25	227.2	2.3	104245	12	ADL34706	Adl34706 Human PPA
C	26	227.2	2.3	170245	12	ADP13586	Adp13586 Renal cel
C	27	225	2.2	25541	4	AAK76168	AAK76168 Human imm
C	28	225	2.2	32248	4	AAK28368	Aas28368 Genomic s
C	29	225	2.2	32248	10	ADG41564	Adg41564 Human res
C	30	225	2.2	32248	11	ADI97338	Adi97338 Human res
C	31	225	2.2	34435	4	AAK76172	AAK76172 Human imm
C	32	224.4	2.2	7815	13	ADS34410	Ads34410 POSH prot
C	33	224.4	2.2	78925	3	AAC89888	Aac89888 Human FN
C	34	224.2	2.2	298	3	AAK07753	Aac07753 Human sec
C	35	224	2.2	66804	6	ABK87050	Abk87050 Human tra
C	36	224	2.2	66804	10	ADG88330	Adg88330 Human tra
C	37	223.6	2.2	3116	4	AAK90031	AAK90031 Human dig
C	38	223.6	2.2	3116	5	AAS39688	Aas39688 Genomic s
C	39	223.6	2.2	3116	9	ADB32648	Abd32648 Human nov
C	40	223.6	2.2	32221	4	ABK42519	Abk42519 Genomic s
C	41	223.6	2.2	32221	9	ADB60675	Adb60675 Connectiv
C	42	223.6	2.2	96594	10	ADC85476	Adc85476 Human Msf
C	43	223.6	2.2	96595	9	ADA02996	Ada02996 Human Msf
C	44	223.6	2.2	96595	10	ADB72734	Adb72734 Human Msf
C	45	223.6	2.2	96595	12	ADM74591	Adm74591 Human car

ALIGNMENTS

RESULT 1

ACN44986
ID ACN44986 standard; DNA; 36022 BP.

AC ACN44986;

DT 18-NOV-2004 (first entry)

DE Human genomic sequence HCG37127.

XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.

XX Homo sapiens.

XX WO2003073826-A2.

XX 12-SEP-2003.

XX 28-FEB-2003; 2003WO-US006235.

XX 01-MAR-2002; 2002US-00087192.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW;

XX WPI; 2003-328604/31.

PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma comprises a nucleotide sequence.

PS Claim 1; SEQ ID NO 1708; Opp; English.

CC The present invention relates to novel DNA and protein sequences which are associated with carcinomas. The sequences are useful for: (i) for screening drug candidates; (ii) for screening of bioactive agent capable of binding to Carcinoma Associated Protein (CAP); (iii) for screening of a bioactive agent capable of modulating the activity of CAP; (iv) for evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining Carcinoma Associated (CA) gene copy number. In addition, the

CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: this patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published
XX
SQ Sequence 36022 BP; 9645 A; 7326 C; 8145 G; 10906 T; 0 U; 0 Other;
Query Match 99.8%; Score 9980.6; DB 11; Length 36022;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 9996; Conservative 0; Mismatches 4; Indels 3; Gaps 1;
Qy 1 CCATCATATTTCTATTTTTTTGGGCGAGAGGGAGACTTCTCTGTGGCCAGGCTGG 60
Db 7528 CCATCATATTTCTATTTTTTTGGGCGAGAGGGAGACTTCTCTGTGGCCAGGCTGG 7587
Qy 61 ACCAGTGGCGCATCTTGCTCAGTCGCAACCTCCACCTCTCGGGTCAAGTGATCCCAA 120
Db 7588 ACCAGTGGCGCATCTTGCTCAGTCGCAACCTCCACCTCTCGGGTCAAGTGATCCCAA 7647
Qy 121 ATAGCTGGGATACAGGTGTGTATTTACCATGCCAGCTAAATTTTGTATTTTACAGAT 180
Db 7648 ATAGCTGGGATACAGGTGTGTATTTACCATGCCAGCTAAATTTTGTATTTTACAGAT 7707
Qy 181 AAGGGTTTCACATGTTGGCCAGGCTGGTCTCCAACTCCTGGCTCATGTGATCCACC 240
Db 7708 AAGGGTTTCACATGTTGGCCAGGCTGGTCTCCAACTCCTGGCTCATGTGATCCACC 7767
Qy 241 ACTTCGGCTTCCCAAGCATTTGGGATAGGTGTGAGCCACTATACCGTCTCACATC 300
Db 7768 ACTTCGGCTTCCCAAGCATTTGGGATAGGTGTGAGCCACTATACCGTCTCACATC 7827
Qy 301 ATATTTCTAATCCGAGACTGTAGAGCTGGTCTCTTTTCTAAGGATGTCACTAGAG 360
Db 7828 ATATTTCTAATCCGAGACTGTAGAGCTGGTCTCTTTTCTAAGGATGTCACTAGAG 7887
Qy 361 AAGTGGAGTTCGCCAAAATTCACGTTTACGTTATTAGTCAAGTTCATAAATACAGTAAT 420
Db 7888 AAGTGGAGTTCGCCAAAATTCACGTTTACGTTATTAGTCAAGTTCATAAATACAGTAAT 7947
Qy 421 AATGTGAGAGTGCATAGGAGCTAACTTGGTTTTTTTTTTTTTTTTTTTTTTTTTTTCAAT 480
Db 7948 AATGTGAGAGTGCATAGGAGCTAACTTGGTTTTTTTTTTTTTTTTTTTTTTTTTTTCAAT 8007
Qy 481 TCTCACTGAACCTTGAATTTGCTTAATAGGACATTAAGGACATTAAGGACATTAAGGAC 540
Db 8008 TCTCACTGAACCTTGAATTTGCTTAATAGGACATTAAGGACATTAAGGACATTAAGGAC 8067
Qy 541 TATTGCTATTGCGCACTATTGATTTTTTAAAAAATAAGCGTATTTTAGCATCTAAAAGT 600
Db 8068 TATTGCTATTGCGCACTATTGATTTTTTAAAAAATAAGCGTATTTTAGCATCTAAAAGT 8127
Qy 601 AGGAAGGACCTCAATAAATAGTCTTTGTTCTTGCCAGGAAACACAGCGTTGTAGAA 660
Db 8128 AGGAAGGACCTCAATAAATAGTCTTTGTTCTTGCCAGGAAACACAGCGTTGTAGAA 8187
Qy 661 TTTGATNACTGTTTTCTAGGCTATGCTGTATTTTCAAGTTAAACCTTGGCGGACGC 720
Db 8188 TTTGATNACTGTTTTCTAGGCTATGCTGTATTTTCAAGTTAAACCTTGGCGGACGC 8247
Qy 721 TAGCATTTAGTAAATCTTTGTAATAGCAATGAACTTAAGCTTCTATGTATAGAAA 780
Db 8248 TAGCATTTAGTAAATCTTTGTAATAGCAATGAACTTAAGCTTCTATGTATAGAAA 8307
Qy 781 CCTAAGTCACTTCACATTTCTGATTTAGCAGAGTAAATGGAATATCTTTTCAATGTAGCT 840
Db 8308 CCTAAGTCACTTCACATTTCTGATTTAGCAGAGTAAATGGAATATCTTTTCAATGTAGCT 8367
Qy 841 CTATCCCGAGACCACAGATATTTGAACTGTAAAGGCCATCTATAGTTTAAACCACTG 900
Db 8368 CTATCCCGAGACCACAGATATTTGAACTGTAAAGGCCATCTATAGTTTAAACCACTG 8427
Qy 901 CGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTTGTGACTA 960
Db

Db 8428 CGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGTTGTGACTA 8487
Qy 961 GAACTCGGGTCTCTGGAGTGTCTTATTTATATCACCAAGCTGTGTACACGCCCATGTGT 1020
Db 8488 GAACTCGGGTCTCTGGAGTGTCTTATTTATATCACCAAGCTGTGTACACGCCCATGTGT 8547
Qy 1021 TGATCTCCATTTGTGATAGCAACAAAGAAAAGACTTTCAGGACATTTCTTTTCTTTTACCCTA 1080
Db 8548 TGATCTCCATTTGTGATAGCAACAAAGAAAAGACTTTCAGGACATTTCTTTTCTTTTACCCTA 8607
Qy 1081 ATCTTTGATCTGCAGTCTTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTATTTCAAAA 1140
Db 8608 ATCTTTGATCTGCAGTCTTTATTTAGAAAAGCTTAATGTTAAAGATCTAGTTTATTTCAAAA 8667
Qy 1141 CTAAGAATAACAAGAGTATGAGAAATTTCTATTTCCGAGTGTAAAGAGGAGAGATTTTCC 1200
Db 8668 CTAAGAATAACAAGAGTATGAGAAATTTCTATTTCCGAGTGTAAAGAGGAGAGATTTTCC 8727
Qy 1201 TTGGCTTCTCTGAGCCTCGAGCCCTTCTTCTGCTTTTAAAGAAAGTAGAGAGGAGGAGAA 1260
Db 8728 TTGGCTTCTCTGAGCCTCGAGCCCTTCTTCTGCTTTTAAAGAAAGTAGAGAGGAGGAGAA 8787
Qy 1261 AGTAAAGTATGCTTTTGTTTTTTAAAGTTACTTTGCTCGGAGTAGTTTTCATGCCCTTTTG 1320
Db 8788 AGTAAAGTATGCTTTTGTTTTTTAAAGTTACTTTGCTCGGAGTAGTTTTCATGCCCTTTTG 8847
Qy 1321 GTTTTCTTGGGTGGAAATTAACGTAACTTAAGTTTTTAAAGTGTGGACATTTTAAAAACAA 1380
Db 8848 GTTTTCTTGGGTGGAAATTAACGTAACTTAAGTTTTTAAAGTGTGGACATTTTAAAAACAA 8907
Qy 1381 TGCTTATCCAATGTTTGCATAAAGCAGAGGGTATTGCTTTTAAAGTTAAATTTCTTCTC 1440
Db 8908 TGCTTATCCAATGTTTGCATAAAGCAGAGGGTATTGCTTTTAAAGTAAATTTCTTCTC 8967
Qy 1441 CAGGAGTGAATAATTAGCTTTCTAAACCAGAACGACAGAGCTAAATAAAGTAAATTTTCCAC 1500
Db 8968 CAGGAGTGAATAATTAGCTTTCTAAACCAGAACGACAGAGCTAAATAAAGTAAATTTTCCAC 9027
Qy 1501 CTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAATGAGAGGCCCATTTTCTGATGA 1560
Db 9028 CTGGCCAGTGCATGATGTGAAAGGTAGATTAAAAAATGAGAGGCCCATTTTCTGATGA 9087
Qy 1561 AAGACTAAGCCATGTTGAAACAGCCCTCTTGAGGATTTTATTTTAAATCTATACATTCAC 1620
Db 9088 AAGACTAAGCCATGTTGAAACAGCCCTCTTGAGGATTTTATTTTAAATCTATACATTCAC 9147
Qy 1621 AAAGAGCTTTGTGTATGCTTTTCCCTATTTGTTTGGACTAGGAAGCCCAAGT 1680
Db 9148 AAAGAGCTTTGTGTATGCTTTTCCCTATTTGTTTGGACTAGGAAGCCCAAGT 9207
Qy 1681 GCTTGTGAAGCCAGAAAAGTGTGAAAGCAAGCTGGGATTTGAAAGTGGATTTGAGTT 1740
Db 9208 GCTTGTGAAGCCAGAAAAGTGTGAAAGCAAGCTGGGATTTGAAAGTGGATTTGAGTT 9267
Qy 1741 TCGAATATTCAGTGAACCAAAATATATCAGGGTTCCTTGGCCAAAGATGAGTGACCATTC 1800
Db 9268 TCGAATATTCAGTGAACCAAAATATATCAGGGTTCCTTGGCCAAAGATGAGTGACCATTC 9327
Qy 1801 TGAGGTGTTAAGTATTTCTTGAATGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 1860
Db 9328 TGAGGTGTTAAGTATTTCTTGAATGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 9387
Qy 1861 ATTTTGTGACCTCTGTATGTGCAAAAATCTCTAAGGGGGTGTTTGGGCACCTTAGATTTCT 1920
Db 9388 ATTTTGTGACCTCTGTATGTGCAAAAATCTCTAAGGGGGTGTTTGGGCACCTTAGATTTCT 9447
Qy 1921 TGGATGTCAGATTTGTTGTATATGAAACAAAATTTTAAATTTGTTTGTATACACTGGAATTT 1980
Db 9448 TGGATGTCAGATTTGTTGTATATGAAACAAAATTTTAAATTTGTTTGTATACACTGGAATTT 9507
Qy 1981 AAAATAGTTTACTAAAGTGTTTTAAATTTTTTTCATCTTAATTTTTCACAGTCTTATAGTCT 2040
Db 9508 AAAATAGTTTACTAAAGTGTTTTAAATTTTTTTCATCTTAATTTTTCACAGTCTTATAGTCT 9567

QY 2041 TTAGATTTAGGAGGCTGTGTGATGGCATCCACATGTGCAATTTTAGTGGCAATTTAAATATG 2100
DB |||||
QY 9568 TTAGATTTAGGAGGCTGTGTGATGGCATCCACATGTGCAATTTTAGTGGCAATTTAAATATG 9627
DB |||||
QY 2101 ATTGAGCTGAATTTAAACAATTTCTGACCTAAACCTTGACATTTTAAAGTTCGGTAA 2160
DB |||||
QY 9628 ATTGAGCTGAATTTAAACAATTTCTGACCTAAACCTTGACATTTTAAAGTTCGGTAA 9687
DB |||||
QY 2161 AGCACTGATTTAAACTGGATTTTAACTGGATGAAATTTCTGATTTTAAAGTTCAGTACT 2220
DB |||||
QY 9688 AGCACTGATTTAAACTGGATTTTAACTGGATGAAATTTCTGATTTTAAAGTTCAGTACT 9747
DB |||||
QY 2221 GGATAAAATGCCAATGATTTTAACTAAAGCAAGCTTTTAAAGGATGCCCTATATATAGT 2280
DB |||||
QY 9748 GGATAAAATGCCAATGATTTTAACTAAAGCAAGCTTTTAAAGGATGCCCTATATATAGT 9807
DB |||||
QY 2281 TAAAGTGAAGCAATTTGAATTTAGGTACTTCTCTGCTGCGTGGTGAAGACCGTATGACTC 2340
DB |||||
QY 9808 TAAAGTGAAGCAATTTGAATTTAGGTACTTCTCTGCTGCGTGGTGAAGACCGTATGACTC 9867
DB |||||
QY 2341 ACCCACACGAGCTTCTCTCGCTCTGAGTGTAGCTAAACCGTTCGTGTTTTTTTCTCTCT 2400
DB |||||
QY 9868 ACCCACACGAGCTTCTCTCGCTCTGAGTGTAGCTAAACCGTTCGTGTTTTTTTCTCTCT 9927
DB |||||
QY 2401 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTTGACTTCTGCTCAATTTGAAACA 2460
DB |||||
QY 9928 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATTTGACTTCTGCTCAATTTGAAACA 9987
DB |||||
QY 2461 CTCAATTTGAAGCAAGAGAACTAATGCTTTGCTGGTGTGATTCATATTTGAATTCAGGC 2520
DB |||||
QY 9988 CTCAATTTGAAGCAAGAGAACTAATGCTTTGCTGGTGTGATTCATATTTGAATTCAGGC 10047
DB |||||
QY 2521 ATTGGCAACCTGTATGCTTGTGTTGGTGAAGAACCGAGTGACACCATCACTGAGCTTC 2580
DB |||||
QY 10048 ATTGGCAACCTGTATGCTTGTGTTGGTGAAGAACCGAGTGACACCATCACTGAGCTTC 10107
DB |||||
QY 2581 TAAAGTTCGAAGAAGTTAGAGGACTATACATTTTCTTTTGAATTTTATATAATAATTT 2640
DB |||||
QY 10108 TAAAGTTCGAAGAAGTTAGAGGACTATACATTTTCTTTTGAATTTTATATAATAATTT 10167
DB |||||
QY 2641 TGCTCTGGTTTTTGGAAACCCAGGGCTGTTAGAGGGTGTAGTGAACAAGTCTTACAAGTGGC 2700
DB |||||
QY 10168 TGCTCTGGTTTTTGGAAACCCAGGGCTGTTAGAGGGTGTAGTGAACAAGTCTTACAAGTGGC 10227
DB |||||
QY 2701 CTATTTCCAACTCCAGAAATTTGCCAACCGGAATTTGAGATTTATATGCAATCGAAAGTGA 2760
DB |||||
QY 10228 CTATTTCCAACTCCAGAAATTTGCCAACCGGAATTTTGAAGATTTATATGCAATCGAAAGTGA 10287
DB |||||
QY 2761 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGATGGGCCAGAAAGTGAATGGCA 2820
DB |||||
QY 10288 CAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGATGGGCCAGAAAGTGAATGGCA 10347
DB |||||
QY 2821 GCTCTCTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATGAAAGGACCGCTGTTG 2880
DB |||||
QY 10348 GCTCTCTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATGAAAGGACCGCTGTTG 10407
DB |||||
QY 2881 TTCCATTTCCGAGCTACACAAGAAATAATGTCAATCCAAATCGAGGGGTATATGCCCTTGG 2940
DB |||||
QY 10408 TTCCATTTCCGAGCTACACAAGAAATAATGTCAATCCAAATCGAGGGGTATATGCCCTTGG 10467
DB |||||
QY 2941 ATTGCAATGTTCTGACGACGAGCTTCAACATTTAGAAAGACCTTTAATAACAATGCTTAA 3000
DB |||||
QY 10468 ATTGCAATGTTCTGACGACGAGCTTCAACATTTAGAAAGACCTTTAATAACAATGCTTAA 10527
DB |||||
QY 3001 TGCAACACCGGCTACCTCTGTGAAACCAAGAGTCTTTCGGGTTGAAGCAGATATCTCA 3060
DB |||||
QY 10528 TGCAACACCGGCTACCTCTGTGAAACCAAGAGTCTTTCGGGTTGAAGCAGATATCTCA 10587
DB |||||
QY 3061 GTCCGCTTGATATAAAGTCAAGTCGGAACAGAACCTCCCAAGGAAGAAATTCGAAGAAA 3120
DB |||||
QY 10588 GTCCGCTTGATATAAAGTCAAGTCGGAACAGAACCTCCCAAGGAAGAAATTCGAAGAAA 10647
DB |||||

QY 3121 ATGAATTTAGCTGTGAGGTATGTGGCAGACATTTTAGAGTCGCTTTTGTATGTTGAGATCC 3180
DB |||||
QY 10648 ATGAATTTAGCTGTGAGGTATGTGGCAGACATTTTAGAGTCGCTTTTGTATGTTGAGATCC 10707
DB |||||
QY 3181 ACATGAGAACACACAAAGATTTCTTTTCACTTACCGGTGTAAACATGTGCGGAAGAAATTC 3240
DB |||||
QY 10708 ACATGAGAACACACAAAGATTTCTTTTCACTTACCGGTGTAAACATGTGCGGAAGAAATTC 10767
DB |||||
QY 3241 AGGAGCTTGGTTTCTTTAAATAATCACAATCGGACACATATGCAAAATCGGGGCCAGAA 3300
DB |||||
QY 10768 AGGAGCTTGGTTTCTTTAAATAATCACAATCGGACACATATGCAAAATCGGGGCCAGAA 10827
DB |||||
QY 3301 GCAAACTGCAGCAAGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTGCTCCAGGTGC 3360
DB |||||
QY 10828 GCAAACTGCAGCAAGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTGCTCCAGGTGC 10887
DB |||||
QY 3361 AGCGGCGGAGAGCATCTCTCTCTTACAAAATCTGCATGTTGTTGGCTTCTTATTTTC 3420
DB |||||
QY 10888 AGCGGCGGAGAGCATCTCTCTCTTACAAAATCTGCATGTTGTTGGCTTCTTATTTTC 10947
DB |||||
QY 3421 CAAATAAAGAAAGTCTAAATTTGAGCACCGCAAGGTGCACACCAAAAACCTGCTTCGGTA 3480
DB |||||
QY 10948 CAAATAAAGAAAGTCTAAATTTGAGCACCGCAAGGTGCACACCAAAAACCTGCTTCGGTA 11007
DB |||||
QY 3481 CCAGCAGCGCGCAGACAGACTCTCCACAAAGGAGAAATGCCGCTCTCGAGGAGAGACTTCC 3540
DB |||||
QY 11008 CCAGCAGCGCGCAGACAGACTCTCCACAAAGGAGAAATGCCGCTCTCGAGGAGAGACTTCC 11067
DB |||||
QY 3541 TGCAGTTGTTCAACTTTGAGACCAAAATCTCACCTGAAACGGGGAAAGACCTTGTCAAT 3600
DB |||||
QY 11068 TGCAGTTGTTCAACTTTGAGACCAAAATCTCACCTGAAACGGGGAAAGACCTTGTCAAT 11127
DB |||||
QY 3601 GCATCCCTCAGCTCGATCCGTTTCCACCATCTTCAGGCTTTGGCAGCTGGCTTACCAAGGAA 3660
DB |||||
QY 11128 GCATCCCTCAGCTCGATCCGTTTCCACCATCTTCAGGCTTTGGCAGCTGGCTTACCAAGGAA 11187
DB |||||
QY 3661 AAGTTGCCATTTGCCAAGAAAGTGAAGAAATCGGGGCAAGAGGAGCACCGACAAACGACG 3720
DB |||||
QY 11188 AAGTTGCCATTTGCCAAGAAAGTGAAGAAATCGGGGCAAGAGGAGCACCGACAAACGACG 11247
DB |||||
QY 3721 ATTGAGTTCCGAGAAAGAGCTTTGGAGAAACAAATTAAGGGCAGTTTGTGAGGCTCTTCGC 3780
DB |||||
QY 11248 ATTGAGTTCCGAGAAAGAGCTTTGGAGAAACAAATTAAGGGCAGTTTGTGAGGCTCTTCGC 11307
DB |||||
QY 3781 AAGAGAAAGAGAAAGTGCRAACACTCCACAGGGCAAGCGCCCTCCGTGGAACCGGATCCCA 3840
DB |||||
QY 11308 AAGAGAAAGAGAAAGTGCRAACACTCCACAGGGCAAGCGCCCTCCGTGGAACCGGATCCCA 11367
DB |||||
QY 3841 AGTTACCCAGTAGCAGAGAGAGCCCACTCACTGCTCCGAGTCCGGCAAAAGCTTTTCAAGAA 3900
DB |||||
QY 11368 AGTTACCCAGTAGCAGAGAGAGCCCACTCACTGCTCCGAGTCCGGCAAAAGCTTTTCAAGAA 11427
DB |||||
QY 3901 CCTACCAACAGCTGTGTTTGCATCTCCAGGGTCCAAAGAGGACCCGAGGCGCGGCGCG 3960
DB |||||
QY 11428 CCTACCAACAGCTGTGTTTGCATCTCCAGGGTCCAAAGAGGACCCGAGGCGCGGCGCG 11487
DB |||||
QY 3961 AGTCCGCCACCACTGTGTGGAACGGGAGGACCGGGGACGTGTTCTCTGACCTCGCG 4020
DB |||||
QY 11488 AGTCCGCCACCACTGTGTGGAACGGGAGGACCGGGGACGTGTTCTCTGACCTCGCG 11547
DB |||||
QY 4021 CCCCTCTGATGAAATAATGAGCCGTGATCGAGGGGAAAGCTGTTCTGAAAGACGGATCTG 4080
DB |||||
QY 11548 CCCCTCTGATGAAATAATGAGCCGTGATCGAGGGGAAAGCTGTTCTGAAAGACGGATCTG 11607
DB |||||
QY 4081 AGGATGGGCTTCCGAGAGAAATCCATCTGGGTAAAGTGTGCCCTGTCTCCGTCCCGTGTGT 4140
DB |||||
QY 11608 AGGATGGGCTTCCGAGAGAAATCCATCTGGGTAAAGTGTGCCCTGTCTCCGTCCCGTGTGT 11667
DB |||||
QY 4141 TCCGCTGTGTGTGTCTCCCGTCTCCCGTCTCCCGTCTCTATTTCCCATCTCCAGACAAACG 4200
DB |||||
QY 11668 TCCGCTGTGTGTGTCTCCCGTCTCCCGTCTCCCGTCTCTATTTCCCATCTCCAGACAAACG 11727
DB |||||
QY 4201 TGGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAAGTCCAGGCTCTTTTTTGGTATCACTCT 4260
DB |||||

[illegible]

QY	6418	CCTATTGTCCTGCGAAGAACTAAAGCCAAAACA	6477	QY	7498	GTGCAATTTTCGGCTCACTGCGAGCTCCACCTC	7557
DB	13948	CCTATTGTCCTGCGAAGAACTAAAGCCAAAACA	14007	DB	15028	GTGCAATTTTCGGCTCACTGCGAGCTCCACCTC	15087
QY	6478	CAGATCTCTTTCCGCCACATGAGTGTGTTGGACA	6537	QY	7558	CCTCCCGAGTAGTGGGATTTACAGGTACCTGCGC	7617
DB	14008	CAGATCTCTTTCCGCCACATGAGTGTGTTGGACA	14067	DB	15088	CCTCCCGAGTAGTGGGATTTACAGGTACCTGCGC	15147
QY	6538	CTAAATTTTGACAGACTTTCCTAGGTGCTTTGCC	6597	QY	7618	TAGTAGAGATGGGGTTTTCACCATGCTGGCCAG	7677
DB	14068	CTAAATTTTGACAGACTTTCCTAGGTGCTTTGCC	14127	DB	15148	TAGTAGAGATGGGGTTTTCACCATGCTGGCCAG	15207
QY	6598	GATTAAAGTTGCTTGTGCTGCCCTCTAGTGTGCT	6657	QY	7678	ATCCGCTCACCTTTGGCCTCCCATAGTGTGGCT	7737
DB	14128	GATTAAAGTTGCTTGTGCTGCCCTCTAGTGTGCT	14187	DB	15208	ATCCGCTCACCTTTGGCCTCCCATAGTGTGGCT	15267
QY	6658	TATTTTTCCTCGTGGTGGGAAGTTGACGGTTTG	6717	QY	7738	AGCCACGGCCCGGACAAAGTTTCATTTGTTTAT	7797
DB	14188	TATTTTTCCTCGTGGTGGGAAGTTGACGGTTTG	14247	DB	15268	AGCCACGGCCCGGACAAAGTTTCATTTGTTTAT	15327
QY	6718	TTCTGAAGAAGATAATTTTCCCGCCAGTATGATG	6777	QY	7798	ATCTTATTTAAAGCTACAGTATTTTAAATGCTC	7857
DB	14248	TTCTGAAGAAGATAATTTTCCCGCCAGTATGATG	14307	DB	15328	ATCTTATTTAAAGCTACAGTATTTTAAATGCTC	15387
QY	6778	GCCTGCTCAGAGACACTTGAGAACCGGAAGCTG	6837	QY	7858	GAAATGAGAACTCTATTTAGTAGTCTTGAGATT	7917
DB	14308	GCCTGCTCAGAGACACTTGAGAACCGGAAGCTG	14367	DB	15388	GAAATGAGAACTCTATTTAGTAGTCTTGAGATT	15447
QY	6838	TTTCTTGTAATTAAGGCAACGGAAGAACTGAAT	6897	QY	7918	GGAGGCTGCTAGATTGGAATTTTCATCTCTTCC	7977
DB	14368	TTTCTTGTAATTAAGGCAACGGAAGAACTGAAT	14427	DB	15448	GGAGGCTGCTAGATTGGAATTTTCATCTCTTCC	15507
QY	6898	GAAAAAACAACAAATCATGTTATAACACTCTAA	6957	QY	7978	AGTTATTTAAACCTTTTGTGCTTTTGTGCTTTT	8037
DB	14428	GAAAAAACAACAAATCATGTTATAACACTCTAA	14487	DB	15508	AGTTATTTAAACCTTTTGTGCTTTTGTGCTTTT	15567
QY	6958	ATTGCTGCTGCTAGCCGTTTGTGTTCTAACCCG	7017	QY	8038	TTCCCTAGGCTGTTTAGAAGATTAAATAAGTTA	8097
DB	14488	ATTGCTGCTGCTAGCCGTTTGTGTTCTAACCCG	14547	DB	15568	TTCCCTAGGCTGTTTAGAAGATTAAATAAGTTA	15627
QY	7018	GCCTTCCAAAATTTTATGTCGAAGATAATTTAA	7077	QY	8098	GAAATAGGCAATTCATAATTTTCAAAATATTCA	8157
DB	14548	GCCTTCCAAAATTTTATGTCGAAGATAATTTAA	14607	DB	15628	GAAATAGGCAATTCATAATTTTCAAAATATTCA	15687
QY	7078	AGCTACCTCTTATATAAATAATTTTGTGTTTCT	7137	QY	8158	GAAATCCTATGATAGTAGATCGAAGAACAAAG	8217
DB	14608	AGCTACCTCTTATATAAATAATTTTGTGTTTCT	14667	DB	15688	GAAATCCTATGATAGTAGATCGAAGAACAAAG	15747
QY	7138	AAAAATAAACATCTTACATCTTCAAGAGAGTGT	7197	QY	8218	TAAATAGATATTTGATCTATTTTTCAGTGTCT	8277
DB	14668	AAAAATAAACATCTTACATCTTCAAGAGAGTGT	14727	DB	15748	TAAATAGATATTTGATCTATTTTTCAGTGTCT	15807
QY	7198	AATATTACCTCAATATTCATCTCAGAACGATAC	7257	QY	8278	CTTGCTTTAGGTGAAAAACCATACAAATGTGAA	8337
DB	14728	AATATTACCTCAATATTCATCTCAGAACGATAC	14787	DB	15808	CTTGCTTTAGGTGAAAAACCATACAAATGTGAA	15867
QY	7258	CCATGCAATAGTTAAATTAAGTATGTTATCTAAT	7317	QY	8338	ACATCTCTGAGGTATCCTTTGGAGAGACATCAC	8397
DB	14788	CCATGCAATAGTTAAATTAAGTATGTTATCTAAT	14847	DB	15868	ACATCTCTGAGGTATCCTTTGGAGAGACATCAC	15927
QY	7318	TGCGAGATCCTTGAGCTAGCTTAAAGGATAAATA	7377	QY	8398	GAAATCAAGAACGATGGTAAAAATCAGGACACT	8457
DB	14848	TGCGAGATCCTTGAGCTAGCTTAAAGGATAAATA	14907	DB	15928	GAAATCAAGAACGATGGTAAAAATCAGGACACT	15987
QY	7378	TTTATTTCTTTAGTGCCATGTAATCTGAGCCATG	7437	QY	8458	GCGCAAAACCAAAAAATTTTGAAGAAGATTTT	8517
DB	14908	TTTATTTCTTTAGTGCCATGTAATCTGAGCCATG	14967	DB	15988	GCGCAAAACCAAAAAATTTTGAAGAAGATTTT	16047
QY	7438	TCTTTCTTTTGTGTTTTGAGACGGAGTCTGCTGT	7497	QY	8518	CCTGCAAGCAGCTTAAAGGAGATGCTTCTGTTT	8577
DB	14968	TCTTTCTTTTGTGTTTTGAGACGGAGTCTTGTCT	15027	DB	16048	CCTGCAAGCAGCTTAAAGGAGATGCTTCTGTTT	16107
				QY	8578	CTCTCACCAGCACACAAAGATACTCAGGATTTT	8637

Db	16108	CTCTCACCAGCACACAAAGATACCTAGGATTTCCATATAAATGACGTGATGACAGTGC	16167	17188	ACATCACTGTTACCCGACGACTGTGTGTATCCGTGCGAGCGCTGCCTCCCAACCAAGG	17247
Qy	8638	GATAAGTGAATATAAAACCTTACCCTGCTTACCTGGACCTGTTAAAAAGAGATCAGCA	8697	9718	TTCTGTAGCTCCAGCGAGGTGATTTCTCCAAATGTGTGACTGTTTCAGAAAGCCCTATGGT	9777
Db	16168	GATAAGTGAATATAAAACCTTACCCTGCTTACCTGGACCTGTTAAAAAGAGATCAGCA	16227	17248	TTCTGTAGCTCCAGCGAGGTGATTTCTCCAAATGTGTGACTGTTTCAGAAAGCCCTATGGT	17307
Qy	8698	GTTGAAACTCAGGCAATTAACCTCATCTGTAGAACCAAGGCGGATGTTACTCTCTCGG	8757	9778	GGCTCCGGCCACATTTTACACTTGTGTGCTCTGCTGTAGTCCAGCATCCAGCTCCGACGTTA	9837
Db	16228	GTTGAAACTCAGGCAATTAACCTCATCTGTAGAACCAAGGCGGATGTTACTCTCTCGG	16287	17308	GGCTCCGGCCACATTTTACACTTGTGTGCTCTGCTGTAGTCCAGCATCCAGCTCCGACGTTA	17367
Qy	8758	GATGCGAGTACCACCATTAACCTTGAAGTTAGCCCCCAAGAGAAACGAGAGCCGCA	8817	9838	GAAGTATTGATGAGGGGCGCTCGTGTGTTAAATGGCTGCCTACAGTGATTAATAGCTAAT	9897
Db	16288	GATGCGAGTACCACCATTAACCTTGAAGTTAGCCCCCAAGAGAAACGAGAGCCGCA	16347	17368	GAAGTATTGATGAGGGGCGCTCGTGTGTTAAATGGCTGCCTACAGTGATTAATAGCTAAT	17427
Qy	8818	GCTGATCGAGATACAGGCCAAGTGTGGATTGTCTACGAAAAACCTTTAAATTTATCCGTG	8877	9898	CCAGGCATTTCAAGTGGAGATGGTACCCTCCCAAGGGTGGGGGTAGGCGCCAGAAAT	9957
Db	16348	GCTGATCGAGATACAGGCCAAGTGTGGATTGTCTACGAAAAACCTTTAAATTTATCCGTG	16407	17428	CCAGGCATTTCAAGTGGAGATGGTACCCTCCCAAGGGTGGGGGTAGGCGCCAGAAAT	17487
Qy	8878	GGGGCTCTTCAATTTGCCGGCAATTTCTTTGAGTAAAGTTTGAATCCCAAGTATCACC	8937	9958	TCCTTGGGGGTCCACAGAGAGAAAGCAATTCCTTAGATAGCGCAGTGG	10000
Db	16408	GGGGCTCTTCAATTTGCCGGCAATTTCTTTGAGTAAAGTTTGAATCCCAAGTATCACC	16467	17488	TCCTTGGGGGTCCACAGAGAGAAAGCAATTCCTTAGATAGCGCAGTGG	17530
Qy	8938	TGTCCATTTTGTACCTTCAAGACATTTTATCCAGAAAGTTTAAATGATGCACCAGAGACTG	8997	RESULT 2		
Db	16468	TGTCCATTTTGTACCTTCAAGACATTTTATCCAGAAAGTTTAAATGATGCACCAGAGACTG	16527	ID	AAV09023 standard; DNA; 10282 BP.	
Qy	8998	GAGCATAAATAAATCTCTGACGTTTATATAAACTGTGCAAAACCAAGTCTCTGTAGAAGT	9057	XX	AAV09023;	
Db	16528	GAGCATAAATAAATCTCTGACGTTTATATAAACTGTGCAAAACCAAGTCTCTGTAGAAGT	16587	XX	21-JUL-1998 (first entry)	
Qy	9058	CSACGTACCGATGCGCGCAGCGTTGCTGGGAAAAGATGTGCTCCCTCCCTAGTTTC	9117	XX	Homo sapiens 20q13 amplicon ZABC-1 genomic sequence.	
Db	16588	CGACGTACCGATGCGCGCAGCGTTGCTGGGAAAAGATGTGCTCCCTCCCTAGTTTC	16647	XX	20q13 amplicon; chromosome 20; tumour; detection; ZABC-1 gene;	
Qy	9118	TCGTAACCCCAAGCTCTCTTTCCCGGCGAGTCCAAATCCCTGCATCTCGAAG	9177	KW	chromosomal abnormalities; probe; gene therapy; antisense inhibition;	
Db	16648	TGTAAACCCCAAGCTCTCTTTCCCGGCGAGTCCAAATCCCTGCATCTCGAAG	16707	KW	treatment; age-related macular degeneration; retinitis pigmentation;	
Qy	9178	GGGAAGCAGAGCCCTCTGGGCCAGGCAAGGCCCTCTGACTTCAGGGATAGACTCTAGC	9237	XX	Leber's congenital amaurosis; zinc finger amplified in breast cancer; ds.	
Db	16708	GGGAAGCAGAGCCCTCTGGGCCAGGCAAGGCCCTCTGACTTCAGGGATAGACTCTAGC	16767	OS	Homo sapiens.	
Qy	9238	ACTTTAGCCCCAAGTAACTGAAGTCCACAGACACAGCAGCAATGTGGGGTCCAAGGG	9297	XX	WO9802539-A1.	
Db	16768	ACTTTAGCCCCAAGTAACTGAAGTCCACAGACACAGCAGCAATGTGGGGTCCAAGGG	16827	PD	22-JAN-1998.	
Qy	9298	GCGGCACCCAGGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCCCTGACCG	9357	PF	15-JUL-1997; 97WO-US012343.	
Db	16828	GCGGCACCCAGGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCCCTGACCG	16887	PR	15-JUL-1996; 96US-00680395.	
Qy	9358	GATAAGACAAAAGACCCGAGACAAAATTTGAAACCTTTCCAGTAGCTCTTTCTCAGCCC	9417	PR	16-OCT-1996; 96US-00731499.	
Db	16888	GATAAGACAAAAGACCCGAGACAAAATTTGAAACCTTTCCAGTAGCTCTTTCTCAGCCC	16947	PR	17-JAN-1997; 97US-00785532.	
Qy	9418	ACCCTCGGAGAGTAAATCATATGTTTCCATCGACTACCCCGCCCAAGAACGAGAGCCCG	9477	XX	(REGC) UNIV CALIFORNIA.	
Db	16948	ACCCTCGGAGAGTAAATCATATGTTTCCATCGACTACCCCGCCCAAGAACGAGAGCCCG	17007	XX	Gray JW, Collins CC, Hwang S, Godfrey T, Kowbel D, Rommens J;	
Qy	9478	TGGGCACCTCCGGGAAGAGACTATTCTGTAATCGGAGTGCAGCAATATCTCAGCAGAA	9537	DR	WPI; 1998-110587/10.	
Db	17008	TGGGCACCTCCGGGAAGAGACTATTCTGTAATCGGAGTGCAGCAATATCTCAGCAGAA	17067	DR	P-PSDB; AAW23975.	
Qy	9538	TTTGTGAGCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCCAG	9597	PT	New sequences from the 20q13 amplicon - used for detecting chromosomal	
Db	17068	TTTGTGAGCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCCAG	17127	XX	abnormalities, particularly tumours, and for developing products for	
Qy	9598	CCCGGGCCAAATACAGAAGAGCTATGACCTTCCCAAGTACCATATGTTAGAGGCATC	9657	PS	treating diseases.	
Db	17128	CCCGGGCCAAATACAGAAGAGCTATGACCTTCCCAAGTACCATATGTTAGAGGCATC	17187	XX	Claim 1; Page 64-67; 91pp; English.	
Qy	9658	ACATCACTGTTACCGCAGGACTGTGTGTATCCGTGCGAGCGCTGCCTCCCAACCAAGG	9717	CC	The sequence is that of the genomic sequence of ZABC-1 (zinc finger	
				CC	amplified in breast cancer). It maps to the core of the 20q13.2 amplicon	
				CC	and is overexpressed in primary tumours and breast cancer cell lines	
				CC	having 20q13.2 amplification. The exact coding region for the genomic	
				CC	sequence is not given. The sequence can be used as a probe for the	
				CC	detection of chromosomal abnormalities at 20q13. It and other sequences	
				CC	isolated from the 20q13 amplicon are consistently amplified in primary	
				CC	tumours. These sequences are useful as probes or as probe targets for	
				CC	monitoring the relative copy number of corresponding sequences from a	
				CC	biological sample such as tumour cells. The sequences can also be used in	

CC therapeutic applications for modulating the expression of the endogenous
CC gene or the activity of the gene product. Examples of therapeutic
CC approaches include antisense inhibition of gene expression, gene therapy,
CC and monoclonal antibodies that specifically bind the gene products. The
CC products can also be used in the treatment of other diseases, e.g. age-
CC related macular degeneration, Leber's congenital amaurosis and retinitis
XX pigmentata

CC Sequence 10282 BP; 2820 A; 2222 C; 2191 G; 3047 T; 0 U; 2 Other;
Query Match 96.58; Score 9651.6; DB 2; Length 10282;
Best Local Similarity 98.58; Pred. No. 0;
Matches 9845; Conservative 2; Mismatches 71; Indels 82; Gaps 7;
QY 1 CCATCATATTTCTTTATTTTGGGGGAGAGGGAGACTTGTCTGTGCTGCCAGGCTGG 60
DB 1 CCATCATATTTCTTTATTTTTCGGCGGAGAGGGAGACTTGTCTGTGCTGCCAGGCTGG 60
QY 61 ACCAGTGGTGGATCTTGGGTCTCATCTGCAACCTCCACTCTCTGGGTTCAAGTGAATCCCAA 120
DB 61 ACCAGTGGTGGATCTTGGGTCTCATCTGCAACCTCCACTCTCTGGGTTCAAGTGAATCCCAA 120
QY 121 ATAGCTGGGATACAGGTGTATTAACCATGCCAGCTAATTTTCTGATTTTATAGCAGAT 180
DB 121 ATAGCTGGGATACAGGTGTATTAACCATGCCAGCTAATTTTCTGATTTTATAGCAGAT 180
QY 181 AAGGGGTTTCCATGTTGGCCAGGCTGGTCTCCAACTCTCTGGGCTCATGTGATCCACCC 240
DB 181 AAGGGGTTTCCATGTTGGCCAGGCTGGTCTCCAACTCTCTGGGCTCATGTGATCCACCC 240
QY 241 ACTTGGGCTTCCCAAAGCATTTGGAGTATAGGTGTGAGGCCATATACCCGTCCTCACATC 300
DB 241 ACTTGGGCTTCCCAAAGCATTTGGAGTATAGGTGTGAGGCCATATACCCGTCCTCACATC 300
QY 301 ATATTTCTAATCCGAGACTGTAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 360
DB 301 ATATTTCTAATCCGAGACTGTAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 360
QY 361 AAGTGGAGTTCCTCCAAATTTACAGTTCACGTATTAGTCAAGTTCCTTAAATACAGTAAT 420
DB 361 AAGTGGAGTTCCTCCAAATTTACAGTTCACGTATTAGTCAAGTTCCTTAAATACAGTAAT 420
QY 421 AATGTTGAGAGTGAATAGGGAATACTTGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 480
DB 421 AATGTTGAGAGTGAATAGGGAATACTTGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 480
QY 481 TCTCACTGAACCTTGGATTTTGTAAATAGGACATTTAAAAAATTAAGGATTTTAAATCCTC 540
DB 481 TCTCACTGAACCTTGGATTTTGTAAATAGGACATTTAAAAAATTAAGGATTTTAAATCCTC 540
QY 541 TATTGCTATTGGCACTATTGATTTTAAAAAATAAGGATTTTATAGCATCTAAAGT 600
DB 541 TATTGCTATTGGCACTATTGATTTTAAAAAATAAGGATTTTATAGCATCTAAAGT 599
QY 601 AGGAAGGACTCAATTAATAGTCTTTGTTTGGCCAGGGAATAACAGGTTGTGAGAA 660
DB 600 AGGAAGGACTCAATTAATAGTCTTTGTTTGGCCAGGGAATAACAGGTTGTGAGAA 659
QY 661 TTTGATACTGTTTTTCTAGGATGTGCTGTTATTTAGTTAAACCTTGGGAGCC 720
DB 660 TTTGATACTGTTTTTCTAGGATGTGCTGTTATTTAGTTAAACCTTGGGAGCC 719
QY 721 TAGCATTCAGTAAATCTTTGTAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 780
DB 720 TAGCATTCAGTAAATCTTTGTAATGAATGAATGAATGAATGAATGAATGAATGAATGA 779
QY 781 CCTAAGTCACTTCCATTTCTGATAGCAGAGTAATTTGAATTTCTTTTCAATTTGTAGCT 840
DB 780 CCTAAGTCACTTCCATTTCTGATTTAGCAGAGTAATTTGAATTTCTTTTCAATTTGTAGCT 839
QY 841 CTATCCCGAGAACCAAGAAATATTTGAATTTGAACTGTAAAGGCCATCTATAGTTTAACTG 900
DB 840 CTATCCCGAGAACCAAGAAATATTTGAATTTGAACTGTAAAGGCCATCTATAGTTTAACTG 899

QY 901 CGTTAAATAGATAATAGAAAGATGTGTATGTGCGAGTGACAACTTGAAGGTTGTGACTA 960
DB 900 CGTTAAATAGATAATAGAAAGATGTGTATGTGCGAGTGACAACTTGAAGGTTGTGACTA 959
QY 961 GAACTCGGCTCTGAGAGTGTCTATATATATACACCAAGCTGTGTACACGAGCCATGTGT 1020
DB 960 GAACTCGGCTCTGAGAGTGTCTATATATATACACCAAGCTGTGTACACGAGCCATGTGT 1019
QY 1021 TGATCTCCATTTGTGATAGCAACAAAGAAAGACTTTCAGGACATTTCTTCTTTACCTTA 1080
DB 1020 TGATCTCCATTTGTGATAGCAACAAAGAAAGACTTTCAGGACATTTCTTCTTTACCTTA 1079
QY 1081 ATCTCTTGATCTGAGTCTTTATTTAGAAAGCTTTAATGTAAAGATCTAGTTTATTTCAAAA 1140
DB 1080 ATCTCTTGATCTGAGTCTTTATTTAGAAAGCTTTAATGTAAAGATCTAGTTTATTTCAAAA 1139
QY 1141 CTAAAGATACAAAGGAGTATAGAAATTTCTATTTCGAGGTGTAAAGAGAGAGATTTTTC 1200
DB 1140 CTAAAGATACAAAGGAGTATAGAAATTTCTATTTCGAGGTGTAAAGAGAGAGATTTTTC 1199
QY 1201 TTGGCTTCTCTGAGCCTGAGGCTTCTCTGCTTTTAAAGAAAGTGTAAAGAGAGAGAGGAA 1260
DB 1200 TTGGCTTCTCTGAGCCTGAGGCTTCTCTGCTTTTAAAGAAAGTGTAAAGAGAGAGAGGAA 1259
QY 1261 AGTAAAGTATGCTTTTGTGTTTTTAAAGGTACTTTGCTGGGAGTGTGATGCTGCTTTTG 1320
DB 1260 AGTAAAGTATGCTTTTGTGTTTTTAAAGGTACTTTGCTGGGAGTGTGATGCTGCTTTTG 1319
QY 1321 GTTTTCTTGGTGGAAATTAACCTGACTTAAAGTGTAAAGTGTGAGGACTATTTTAAAAACAA 1380
DB 1320 GTTTTCTTGGTGGAAATTAACCTGACTTAAAGTGTAAAGTGTGAGGACTATTTTAAAAACAA 1379
QY 1381 TGCTATCCAAATTTTGGCCATAAAGCAGAGGCTATTTGCTTTAGAAAGTAAATTTCTCTC 1440
DB 1380 TGCTATCCAAATTTTGGCCATAAAGCAGAGGCTATTTGCTTTAGAAAGTAAATTTCTCTC 1439
QY 1441 CAGGAGTGAATAATAGTCTTAAACCCAGAGCAGAGCTAAATAAGGTAATTTTTCAC 1500
DB 1440 CAGGAGTGAATAATAGTCTTAAACCCAGAGCAGAGCTAAATAAGGTAATTTTTCAC 1499
QY 1501 CTGGCCAGTGCATGATGTGAAAGTGTGAAAGTGTGAAAGTGTGAAAGTGTGAAAGTGTG 1560
DB 1500 CTGGCCAGTGCATGATGTGAAAGTGTGAAAGTGTGAAAGTGTGAAAGTGTGAAAGTGTG 1559
QY 1561 AAGACTAAGCCATTTGAAACAGCCCTGTGAGGATTTTATTTTAAATCTATACATTCAC 1620
DB 1560 AAGACTAAGCCATTTGAAACAGCCCTGTGAGGATTTTATTTTAAATCTATACATTCAC 1619
QY 1621 AAAGGAGCTTTGTGTATGTCTTTTCCCTATTTGTTGTTGGACTAGGAAGCCCAACCCAGT 1680
DB 1620 AAAGGAGCTTTGTGTATGTCTTTTCCCTATTTGTTGTTGGACTAGGAAGCCCAACCCAGT 1679
QY 1681 GCTTTGTTGAAGCAGAAAGTCTGTGAAAGCAAGCTGGGATTTTGAACAGTGGATTTGAGGTT 1740
DB 1680 GCTTTGTTGAAGCAGAAAGTCTGTGAAAGCAAGCTGGGATTTTGAACAGTGGATTTGAGGTT 1739
QY 1741 TCGAATATCCAGTGAACCAAAATATATACAGGTTCCCTGGCCCAAGATGAGTGACCATTC 1800
DB 1740 TCGAATATCCAGTGAACCAAAATATATACAGGTTCCCTGGCCCAAGATGAGTGACCATTC 1799
QY 1801 TCAGGTTGTTAAGTATTTCTTGAATGGGATTTTGAAGAAAGTTTCTGTATTTCTGTGCTC 1860
DB 1800 TCAGGTTGTTAAGTATTTCTTGAATGGGATTTTGAAGAAAGTTTCTGTATTTCTGTGCTC 1859
QY 1861 ATTTTGTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTGTTGGGCACTTAGATTTCT 1920
DB 1860 ATTTTGTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTGTTGGGCACTTAGATTTCT 1919
QY 1921 TCGATGACAGATTTGTTGTATATGAAAACAAATTTTAAATTTTGTGTATACACTGATTT 1980
DB 1920 TCGATGACAGATTTGTTGTATATGAAAACAAATTTTAAATTTGTTTGTATACACTGATTT 1979

1981 AAAATAGTTTACTAAAGTGTGTTTAAATTTTTCATCTTAATTTTACAGTTCCTTATAGTCT 2040
1980 AAAATAGTTTACTAAAGTGTGTTTAAATTTTTCATCTTAATTTTACAGTTCCTTATAGTCT 2039
2041 TTAGATTTTAGGGAGGCTGTGATGGCATCCACATGTGCATTTTAGTGCAATTTAAATATGT 2100
2040 TTAGATTTTAGGGAGGCTGTGATGGCATCCACATGTGCATTTTAGTGCAATTTAAATATGT 2099
2101 ATTACGTGAATTTAAACAAATTTCTGACCTTAAACCTTGACATTTTATAGTTCGGTAA 2160
2100 ATTACGTGAATTTAAACAAATTTCTGACCTTAAACCTTGACATTTTATAGTTCGGTAA 2159
2161 AGCACTGAATTTAACTGGATTTTAACTGGATGAATTTCTGATTTTAAATAGTGTACTGACT 2220
2160 AGCACTGAATTTAACTGGATTTTAACTGGATGAATTTCTGATTTTAAATAGTGTACTGACT 2219
2221 GGATTAATAATGCCAATGATTTAATTAACAAAGCACGTTTAAACAGATGCCCTTATATATAGT 2280
2220 GGATTAATAATGCCAATGATTTAATTAACAAAGCACGTTTAAACAGATGCCCTTATATATAGT 2279
2281 TAAAGTGAAGCAATGAAATAGTACCTTCTCTGCTGCTGGTGGAAAGACCGTATGACTC 2340
2280 TAAAGTGAAGCAATGAAATAGTACCTTCTCTGCTGCTGGTGGAAAGACCGTATGACTC 2339
2341 ACCCACACGAGCTTCTCTGCTCTGAGTGTAGCTTAACCGTTTCTGTTTCTTCTCTCT 2400
2340 ACCCACACGAGCTTCTCTGCTCTGAGTGTAGCTTAACCGTTTCTGTTTCTTCTCTCT 2399
2401 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATGACTTCTTGTCTCAATTTGAAACA 2460
2400 AGGGTTTGGAAATCCCTTGTCTCCAGGTTGCTGGGATGACTTCTTGTCTCAATTTGAAACA 2459
2461 CTCAATTCATGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 2520
2460 CTCAATTCATGAGACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGC 2519
2521 ATTGGGAACCTGATGCTTCTGTTGTGGAAAGAACCCAGTGCACCATCAGCTTCC 2580
2520 ATTGGGAACCTGATGCTTCTGTTGTGGAAAGAACCCAGTGCACCATCAGCTTCC 2579
2581 TAAAGTTTCGAAGAGTTAGAGGACTATACATCTTCTTTTGAATCTTTTATATATATAT 2640
2580 TAAAGTTTCGAAGAGTTAGAGGACTATACATCTTCTTTTGAATCTTTTATATATATAT 2638
2641 TGCTCTGTTTGGAAACCCAGGCTGTTAGAGGGTGTAGTGAACAAGTCTTCAAGTGGC 2700
2639 TGCTCTGTTTGGAAACCCAGGCTGTTAGAGGGTGTAGTGAACAAGTCTTCAAGTGGC 2698
2701 CTTATTCCAACTCCAGAAATTCGCCAACGGAATTTGAGATTTATGCAATCGAAAGTGA 2760
2699 CTTATTCCAACTCCAGAAATTCGCCAACGGAATTTGAGATTTATGCAATCGAAAGTGA 2758
2761 CAGGAAACATGCAACTCAATCCCTCTTAATGTACATGATGGGCGCAGAGTGTGGCA 2820
2759 CAGGAAACATGCAACTCAATCCCTCTTAATGTACATGATGGGCGCAGAGTGTGGCA 2818
2821 GCTCTCTTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATGAAAGGACCGCTGTG 2880
2819 GCTCTCTTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATGAAAGGACCGCTGTG 2878
2881 TTTCCATTCGAGCTACACAAGAAATAATGTCATCCAAATCAGGGGTATATGCCCTTGG 2940
2879 TTTCCATTCGAGCTACACAAGAAATAATGTCATCCAAATCAGGGGTATATGCCCTTGG 2938
2941 ATTGATGTTTCTGCGCAGACCTTTCACATTTTCAAGACCTTAAATAACATGTCTTAA 3000
2939 ATTGATGTTTCTGCGCAGACCTTTCACATTTTCAAGACCTTAAATAACATGTCTTAA 2998
3001 TGCAACACCGGCTACCTCTCTGTGAACGAGAGTCTTCCGGTGTGAAGCAGATATCTCA 3060
2999 TGCAACACCGGCTACCTCTCTGTGAACGAGAGTCTTCCGGTGTGAAGCAGATATCTCA 3058
3061 GTCCGCTTGATTAAGTCAAGTGCAGAACAGAACCTCCCAAGGAAAGAAATTCAGAGAA 3120

3059 GTCCGCTTGATTAAGTCAAGTGCAGAACAGAACCTCCCAAGGAAAGAAATTCAGAGAA 3118
3121 ATGAATTTAGTGTGAGGTATGTGGCAGACATTTTAGAGTGCCTTTTGTATGTTGATCTC 3180
3119 ATGAATTTAGTGTGAGGTATGTGGCAGACATTTTAGAGTGCCTTTTGTATGTTGATCTC 3178
3181 ACATGAGAAACACACAAAGATTTCTTCACTTAGGGGTGTAAACATGTGCGGAAGAAATTC 3240
3179 ACATGAGAAACACACAAAGATTTCTTCACTTAGGGGTGTAAACATGTGCGGAAGAAATTC 3238
3241 AGGAGCTTTGGTTCTTAAATAATCATGTGCGGACACATAATGGCAAAATCGGGGGCAGAA 3300
3239 AGGAGCTTTGGTTCTTAAATAATCATGTGCGGACACATAATGGCAAAATCGGGGGCAGAA 3298
3301 GCAAACTGCAGCAAGGCTTTGGAGAGTGTTCAGAACACGATCAACGAGGTGCTCCAGGTGC 3360
3299 GCAAACTGCAGCAAGGCTTTGGAGAGTGTTCAGAACACGATCAACGAGGTGCTCCAGGTGC 3358
3361 ACGGGCCGAGAGCATCTCTCTCTTACAAAATCTGCATGGTGTGCTTCTTCTATTTTC 3420
3359 ACGGGCCGAGAGCATCTCTCTCTTACAAAATCTGCATGGTGTGCTTCTTCTATTTTC 3418
3421 CAAATTAAGAAAGTCTAATTTGAGCACCGCAAGTGTGCACCAAAAAAACTGCTTTCGGTA 3480
3419 CAAATTAAGAAAGTCTAATTTGAGCACCGCAAGTGTGCACCAAAAAAACTGCTTTCGGTA 3478
3481 CCAGCAGCGCGCAGACAGACTTCCACAAGGAGGAATGCGTCTCCGAGGAGGACTTCC 3540
3479 CCAGCAGCGCGCAGACAGACTTCCACAAGGAGGAATGCGTCTCCGAGGAGGACTTCC 3538
3541 TGCACTTGTTCMACTTTGAGACCAAAATCTCACCCTGAAACCGGGGAGAGCCCTGTCAGAT 3600
3539 TGCACTTGTTCMACTTTGAGACCAAAATCTCACCCTGAAACCGGGGAGAGCCCTGTCAGAT 3597
3601 GCATCCCTCAGCTCCGATCCGTTACACCTTCCAGGCTTGGCAGCTGGCTACCAAGGAA 3660
3598 GCATCCCTCAGCTCCGATCCGTTACACCTTCCAGGCTTGGCAGCTGGCTACCAAGGAA 3657
3661 AAGTTGCCATTTGCAAGAGTGAAGAAATCGGGGCAAGAGGAGACACCGAACAGCG 3720
3658 AAGTTGCCATTTGCAAGAGTGAAGAAATCGGGGCAAGAGGAGACACCGAACAGCG 3717
3721 ATTCCAGTTCCGAGAGGAGCTTGGAGAAACAAATAAGGGCAGTGTGCGAGCCCTCTCC 3780
3718 ATTCCAGTTCCGAGAGGAGCTTGGAGAAACAAATAAGGGCAGTGTGCGAGCCCTCTCC 3777
3781 AAGGAAAGAGAGTGCACAACTCCACGGGCAAGCGCCCTCCGTTGACGCGGATCCCA 3840
3778 AAGGAAAGAGAGTGCACAACTCCACGGGCAAGCGCCCTCCGTTGACGCGGATCCCA 3837
3841 AGTTACCCAGTAGCAAGGAGAGCCCACTCACTGCTCCGAGTGCAGCAAAAGCTTTTCAGAA 3900
3838 AGTTACCCAGTAGCAAGGAGAGCCCACTCACTGCTCCGAGTGCAGCAAAAGCTTTTCAGAA 3897
3901 CCTACACAGCTGCTTGTGCACTCCAGGTTCCAAGAGAGACCGGAGGCGGCGCG 3960
3898 CCTACACAGCTGCTTGTGCACTCCAGGTTCCAAGAGAGACCGGAGGCGGCGCG 3957
3961 AGTCCGCCCAACATGCTGTGAGCGGAGCGGAGCGGAGCGGAGCTTCTCTGACCTCGCG 4020
3958 AGTCCGCCCAACATGCTGTGAGCGGAGCGGAGCGGAGCGGAGCTTCTCTGACCTCGCG 4017
4021 CCCCTCTGGATGAAATGGAGCCGTGATTCGAGGGGAGGTTGTTCTGAAAGCGGATCTG 4080
4018 CCCCTCTGGATGAAATGGAGCCGTGATTCGAGGGGAGGTTGTTCTGAAAGCGGATCTG 4077
4081 AGGATGGGCTTCCCAAGGAAATCCATCTGGGTAAAGTGCCTGCTCCGTCCTGCTGT 4140
4078 AGGATGGGCTTCCCAAGGAAATCCATCTGGGTAAAGTGCCTGCTCCGTCCTGCTGT 4100
4141 TCCGCTGT 4200

Db 4101 -----CTATTCCTCCATCTCCAGCAACGC 4123
Qy 4201 TGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTGGTATCACTCT 4260
Db 4124 TGCCAGGAATGGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTGGTATCACTCT 4183
Qy 4261 GTGTAGTCAATTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAAATACAGGGTTGA 4320
Db 4184 GTGTAGTCAATTAACCTCTCAGGGCCTTAATTTTCTCATTTCTGTAAATACAGGGTTGA 4243
Qy 4321 GTTAAAGAGTCTCCTCTGCTTCTGAAAAATATATATATATTTTAAACGTGTATCGTTTGC 4380
Db 4244 GTTAAAGAGTCTCCTCTGCTTCTGAAAAATATATATATATTTTAAACGTGTATCGTTTGC 4303
Qy 4381 TCACAAACACACTTTAAAAAATAAATACTTGTGCATCCAGGCCAAATGCACTGCTTCT 4440
Db 4304 TCACAAACACACTTTAAAAAATAAATACTTGTGCATCCAGGCCAAATGCACTGCTTCT 4363
Qy 4441 TAACTGGGGCGATTTTGTCCCAATCAGTATCTGGCAATGTCTGGAGGCAATTTTGGTTGT 4500
Db 4364 TAACTGGGGCGATTTTGTCCCAATCAGTATCTGGCAATGTCTGGAGGCAATTTTGGTTGT 4423
Qy 4501 CATACTGTGTGTGGGTGTGCTGCTGGCATCCAGTGGCAGAGGCCAGGACATGCT 4560
Db 4424 CATACTGTGTGTGGGTGTGCTGCTGGCATCCAGTGGCAGAGGCCAGGACATGCT 4483
Qy 4561 CAGCATGGTACAGTGCACAGACAGCCCCATCATCAAGAAATATCTGGTCCCAATGTC 4620
Db 4484 CAGCATGGTACAGTGCACAGACAGCCCCATCATCAAGAAATATCTGGTCCCAATGTC 4543
Qy 4621 AATAGTTTGAAGCAATGAGAGACCTAGCCCTTCACTTAAAGTTTCTGGGCTTCCCTGATCT 4680
Db 4544 AATAGTTTGAAGCAATGAGAGACCTAGCCCTTCACTTAAAGTTTCTGGGCTTCCCTGATCT 4603
Qy 4681 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAGGTACTGGAGTGATCACTAGAGCCAG 4740
Db 4604 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAGGTACTGGAGTGATCACTAGAGCCAG 4663
Qy 4741 GAATATTATTGGGCGAGCCGTTTGGTGTCTGTCACAAACCTTGTCTTCTGTCTGGCAAG 4800
Db 4664 GAATATTATTGGGCGAGCCGTTTGGTGTCTGTCACAAACCTTGTCTTCTGTCTGGCAAG 4723
Qy 4801 CTAGTATCCATTTATAGGTACTCTCAGGAACCCAAATGATTTGTCTATAAATAACAAGGAAT 4860
Db 4724 CTAGTATCCATTTATAGGTACTCTCAGGAACCCAAATGATTTGTCTATAAATAACAAGGAAT 4783
Qy 4861 GTGAGCACACTGAAGACATTTTAAAGGCTCAATTTGCTCAGCAGAAATTTTCAAGTGTAC 4920
Db 4784 GTGAGCACACTGAAGACATTTTAAAGGCTCAATTTGCTCAGCAGAAATTTTCAAGTGTAC 4843
Qy 4921 TAGTGGCATTTATAGAAAGAGAGGTGATCACTGAAGGATGCTCACATAATATCTCTGA 4980
Db 4844 TAGTGGCATTTATAGAAAGAGAGGTGATCACTGAAGGATGCTCACATAATATCTCTGA 4903
Qy 4981 GGCCTGGTGGGCTTATCTAGGCAAGGATCCACCTGTGTTTGGAGTTGGGCCCATCC 5040
Db 4904 GGCCTGGTGGGCTTATCTAGGCAAGGATCCACCTGTGTTTGGAGTTGGGCCCATCC 4963
Qy 5041 TCACTGTAGCCAGGCTTCTCCTATCAGAGTTAGTATTTTGTGAATAGAGATCTTG 5100
Db 4964 TCACTGTAGCCAGGCTTCTCCTATCAGAGTTAGTATTTTGTGAATAGAGATCTTG 5023
Qy 5101 CTGCTTTAAACAGTGTGAAAGAGCCCTGATGGGCGAGCCGTAATTTGACAGCGGAATCATG 5160
Db 5024 CTGCTTTAAACAGTGTGAAAGAGCCCTGATGGGCGAGCCGTAATTTGACAGCGGAATCATG 5083
Qy 5161 GAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCTTGGTTAAACCAAGTGCC 5220
Db 5084 GAACATGAATCGGTCTTAGGGAAGCATCTGTCAAAGTGGTCTTGGTTAAACCAAGTGCC 5143
Qy 5221 TCCTCCTCTCAGTGTCACTTGAATGTGTGTTGAATTTCTCGGAAACCTGGGTGTATGAG 5280
Db 5144 TCCTCCTCTCAGTGTCACTTGAATGTGTGTTGAATTTCTCGGAAACCTGGGTGTATGAG 5203

Qy 5281 ACCCAGATGAATTTGGCCACACAGATTTGAATTTGACTCTTCTTCACTGCTCTTTCAGCCA 5340
Db 5204 ACCCAGATGAATTTGGCCACACAGATTTGAATTTGACTCTTCTTCACTGCTCTTTCAGCCA 5263
Qy 5341 GTGCCAGTTCCTTTTCTGATCATGTGATTCAGTGAACCTGTAGTCTGTATATCAAAATC 5400
Db 5264 CTGCCAGTTCCTTTTCTGATCATGTGATTCAGTGAACCTGTAGTCTGTATATCAAAATC 5323
Qy 5401 TTTAGAAATGTTTTAGTTTCTGGGACACAGGAAACCCAGCACTTAGCATACTACAAT 5460
Db 5324 TTTAGAAATGTTTTAGTTTCTGGGACACAGGAAACCCAGCACTTAGCATACTACAAT 5383
Qy 5461 CTAAATGCTTAAATGGGATCATATAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAAGTG 5520
Db 5384 CTAAATGCTTAAATGGGATCATATAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAAGTG 5443
Qy 5521 GTTTCGTAGTCCCGTACTGTTCAGAGGGCCCTGTGAGATGCCAGTTCCTCGTGAAG 5580
Db 5444 GTTTCGTAGTCCCGTACTGTTCAGAGGGCCCTGTGAGATGCCAGTTCCTCGTGAAG 5503
Qy 5581 AAATGAAGAGGCCAGTTACCGGTAGTGGTGTGGAAAAATCATGGGCTAGATCATCAGGCAG 5640
Db 5504 AAATGAAGAGGCCAGTTACCGGTAGTGGTGTGGAAAAATCATGGGCTAGATCATCAGGCAG 5563
Qy 5641 GACAGAAATGCTGGCTGTGGGTGGGAGCACCCAGCTTTGGCGTTGAGTTCTGTCTTACC 5700
Db 5564 GACAGAAATGCTGGCTGTGGGTGGGAGCACCCAGCTTTGGCGTTGAGTTCTGTCTTACC 5623
Qy 5701 ACTGCGTGTGTTGTGACCAATTTATGAGTTGTCTTAAACCTTTTCTTGTCTACTATTTCCCTG 5760
Db 5624 ACTGCGTGTGTTGTGACCAATTTATGAGTTGTCTTAAACCTTTTCTTGTCTACTATTTCCCTG 5683
Qy 5761 TTTGCAAAATGGTTCAATTCAGCCCTCTCTTCCACCTCCCAAGCAATTTCAACAGGCCTA 5820
Db 5684 TTTGCAAAATGGTTCAATTCAGCCCTCTCTTCCACCTCCCAAGCAATTTCAACAGGCCTA 5743
Qy 5821 TTTGTAATAAAGATCAAGTCTTTTAAATAATAAATGTAAGTCAAGAGTGATGCTTGA 5880
Db 5744 TTTGTAATAAAGATCAAGTCTTTTAAATAATAAATGTAAGTCAAGAGTGATGCTTGA 5803
Qy 5881 AAGAGCAGGAAACAGGTAGATGTGGAATGTCTATGCTCTTTGTTCTTAAAGAAAGGCATT 5940
Db 5804 AAGAGCAGGAAACAGGTAGATGTGGAATGTCTATGCTCTTTGTTCTTAAAGAAAGGCATT 5863
Qy 5941 TCATAGCTTTTGGATATGACGCAACATACCAATTAATCTGACACATAGTTGGAGTCCG 6000
Db 5864 TCATAGCTTTTGGATATGACGCAACATACCAATTAATCTGACACATAGTTGGAGTCCG 5923
Qy 6001 AAATTCGCAACACGCGCCAGTTATAACCCAGCTAGTTTGGGTATGATTTGTAAGAAATAA 6060
Db 5924 AAATTCGCAACACGCGCCAGTTATAACCCAGCTAGTTTGGGTATGATTTGTAAGAAATAA 5983
Qy 6061 AGCTGSCCATCTCTGATTTTGGGAAATGATTTTCTTAAACTTATATTTCTTAGTAGTCT 6120
Db 5984 AGCTGSCCATCTCTGATTTTGGGAAATGATTTTCTTAAACTTATATTTCTTAGTAGTCT 6043
Qy 6121 AGATTTATCATATTTGACTATCATCTCTGCTTTTAAAGACTTAAAGACTCAAGTAAAT 6180
Db 6044 AGATTTATCATATTTGACTATCATCTCTGCTTTTAAAGACTTAAAGACTCAAGTAAAT 6103
Qy 6181 TTTTCTTCTTCTTTTAGACACTATATAGATCATCAAGGCTGTCTGCTTACAGGTGGAT 6240
Db 6104 TTTTCTTCTTCTTTTAGACACTATATAGATCATCAAGGCTGTCTGCTTACAGGTGGAT 6163
Qy 6241 AGTGATATGATCTACAGTGAGGGGCAATTTTAAACCTTAAACATTCATGTGTTTTGG 6300
Db 6164 AGTGATATGATCTACAGTGAGGGGCAATTTTAAACCTTAAACATTCATGTGTTTTGG 6223
Qy 6301 GGGTGTATTTTAAACGAGCAGCTCTGATTTGTTTTGGAGGCTGTGTGTTTTGAA 6360
Db 6224 GGGTGTATTTTAAACGAGCAGCCTCTGATTTGTTTTGGAGGCTGTGTGTTTTGAA 6283

8439	DB	GC	AACACCACTTAA	GGAGATG	CCCTTC	TGTTTTT	CAGAA	ATGTTCTG	CCCA	CCCCCTCT	CCCTC	8499													
8581	QY	TC	CCAGCACACA	AAAGATAC	TCTCAG	ATATTC	CA	AAAAAT	CAGCTG	ATGACAG	TGCTGAT	8640													
8499	DB	TC	CCACCA	CACAAAGAT	TACTCAG	ATATTC	CA	AAAAAT	CAGCTG	ATGACAG	TGCTGAT	8558													
8641	QY	AA	GTGAATA	AAAAAC	CTAC	CCCTG	CTTA	CCTG	GACCTG	TTAAAA	AGAGATCAG	CAAGTT	8700												
8559	DB	AA	GTGAATA	AAAAAC	CTAC	CCCTG	CTTA	CCTG	GACCTG	TTAAAA	AGAGATCAG	CAAGTT	8618												
8701	QY	GA	AACTCAGG	CAATAA	CTCAT	CTGT	AGAAC	CAAG	CGGATG	TTACT	CTCTC	CCGGAT	8760												
8619	DB	GA	AACTCAGG	CAATAA	CTCAT	CTGT	AGAAC	CAAG	CGGATG	TTACT	CTCTC	CCGGAT	8678												
8761	QY	GG	CAGTAC	CA	CCCA	TAACTTT	GAA	GTAG	CCCA	AAAGAA	GAGAAAC	CGAGACCG	CAAGCT	8820											
8679	DB	GG	CAGTAC	CA	CCCA	TAACTTT	GAA	GTAG	CCCA	AAAGAA	GAGAAAC	CGAGACCG	CAAGCT	8738											
8821	QY	GAC	TGCAGAT	ACAGG	CCAA	GTGTG	GGATG	TG	TCAC	GAAAAA	CTTTAA	ATTTAT	CCGTGGGG	8880											
8739	DB	GAC	TGCAGAT	ACAGG	CCAA	GTGTG	GGATG	TG	TCAC	GAAAAA	CTTTAA	ATTTAT	CCGTGGGG	8798											
8881	QY	GCT	CTTCC	CA	CAAT	TG	CCCGG	CAATTT	CTT	TCAG	TAAAG	TTTGAT	TCCAA	GATFATCA	CTGT	8940									
8799	DB	GCT	CTTCC	CA	CAAT	TG	CCCGG	CAATTT	CTT	TCAG	TAAAG	TTTGAT	TCCAA	GATFATCA	CTGT	8858									
8941	QY	CC	ATTTTGT	TAC	TTCA	AGAC	CA	ATTTT	T	CCAG	AA	GTTTT	AT	TGATG	CA	CCAGAC	TGGAG	9000							
8859	DB	CC	ATTTTGT	TAC	TTCA	AGAC	CA	ATTTT	T	CCAG	AA	GTTTT	AT	TGATG	CA	CCAGAC	TGGAG	8918							
9001	QY	CAT	AAATAC	AAAT	CTCTG	AT	CA	TAAAAA	CTG	CG	AAAC	CAAGTCT	TG	CTT	TAG	AA	GTGCA	9060							
8919	DB	CAT	AAATAC	AAAT	CTCTG	AT	CA	TAAAAA	CTG	CG	AAAC	CAAGTCT	TG	CTT	TAG	AA	GTGCA	8978							
9061	QY	CG	TAC	CGGATG	CCCG	CCAG	CGT	TG	CTGG	AAAA	AGATG	TC	CCCT	CCCT	CTAG	TTCT	GT	9120							
8979	DB	CG	TAC	CGGATG	CCCG	CCAG	CGT	TG	CTGG	AAAA	AGATG	TC	CCCT	CCCT	CTAG	TTCT	GT	9038							
9121	QY	AA	ACCC	AAAG	CCCA	AGTCTG	CT	TTT	CCCGG	CGAGT	CA	AAAT	CCCT	TG	CCAT	CTG	CG	AAAGGG	9180						
9039	DB	AA	ACCC	AAAG	CCCA	AGTCTG	CT	TTT	CCCGG	CGAGT	CA	AAAT	CCCT	TG	CCAT	CTG	CG	AAAGGG	9098						
9181	QY	AAG	CAG	AG	CC	CT	CTG	GGC	CAGG	CA	AG	CC	CT	CT	GAT	TTC	AGG	ATAG	CT	CTAG	CAC	T	9240		
9099	DB	AAG	CAG	AG	CC	CT	CTG	GGC	CAGG	CA	AG	CC	CT	CT	GAT	TTC	AGG	ATAG	CT	CTAG	CAC	T	9158		
9241	QY	TT	AG	CCCA	AGTAA	CTG	AG	ATG	TTT	CT	CA	AAAA	CA	AGT	GT	TTT	CCCT	CTG	GGG	TCC	AA	GGGG	9300		
9159	DB	TT	AG	CCCA	AGTAA	CTG	AG	ATG	TTT	CT	CA	AAAA	CA	AGT	GT	TTT	CCCT	CTG	GGG	TCC	AA	GGGG	9218		
9301	QY	GC	CA	CCAGG	CA	CAG	CAAT	CTG	AG	ATG	TTT	CT	CA	AAAA	CA	AGT	GT	TTT	CCCT	CTG	GGG	TCC	AA	GGGG	9360
9219	DB	GC	CA	CCAGG	CA	CAG	CAAT	CTG	AG	ATG	TTT	CT	CA	AAAA	CA	AGT	GT	TTT	CCCT	CTG	GGG	TCC	AA	GGGG	9278
9361	QY	AAG	CA	AAAA	AGA	CC	CG	AG	CA	AAAA	AT	GAA	AC	CT	TTT	CC	AG	TAG	CT	CT	CT	CAG	CC	CA	9420
9279	DB	AAG	CA	AAAA	AGA	CC	CG	AG	CA	AAAA	AT	GAA	AC	CT	TTT	CC	AG	TAG	CT	CT	CT	CAG	CC	CA	9338
9421	QY	CT	GG	CAG	CA	GTAA	CA	TCA	TG	TT	CG	AT	CA	CC	CG	CC	AA	AG	CA</						

RESULT 3

RESUL 3
AAK81093

AAK81093
ID AAK81093 standard: DNA: 14906 BP.

[illegible]

AC AAK81093;

XX

DT 07-NOV-2001 (first entry)

XX

DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35905.

XX

KW Human; immune

KW cytostatic; c

XX

OS Homo sapiens.

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

PN WO200157182-A1

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

PD 09-AUG-2001.
vv

XX
DE 17-JAN-200317-JAN-2001;
PF
YY

31 - TAN - 2000 -

PR 31-JAN-2000;
PR 04-FEB-2000:

PR 04-FEB-2000;
PR 24-FEB-2000:

PR 02-MAR-2000:
PR 24-FEB-2000:

PR 16-MAR-2000:

17-MAR-2000;
PR

PR 18-APR-2000;

PR 19-MAY-2000;

PR 07-JUN-2000;

PR 28-JUN-2000;

PR 30-JUN-2000;

PR 07-JUL-2000;

PR 07-JUL-2000;

PR 11-JUL-2000;

PR 11-JUL-2000;

PR 14-JUL-2000;

PR 26-JUL-2000;

PR 26-JUL-2000;
DD 14 AUG 2000

PR 14-AUG-2000;
PP 14 AUG 2000

PR 14-AUG-2000;
PR 14-AUG-2000-

QY 3601 GCATCCCTCAGTCGATCCGTTACACCTTTCAGGCTTGCGAGCTGGCTACCAAGGAA 3660
DB 1 GCATCCCTCAGTCGATCCGTTACACCTTTCAGGCTTGCGAGCTGGCTACCAAGGAA 60
QY 3661 AAGTTGCAATTTGCAAGAAGTGAAGAAATCGGGCAAGAGGAGCACGCAACAGACG 3720
DB 61 AAGTTGCAATTTGCAAGAAGTGAAGAAATCGGGCAAGAGGAGCACGCAACAGACG 120
QY 3721 ATTGAGTTCGGAAGAGCTTGGGAAACAAATAAGGGCAGTTGTGCAAGGCTCTCGC 3780
DB 121 ATTGAGTTCGGAAGAGCTTGGGAAACAAATAAGGGCAGTTGTGCAAGGCTCTCGC 180
QY 3781 AAGAGAAGAGAGTGCACAAACATCCCAACGCGGAAGCGCCCTCGTGGAACGCGAATCCCA 3840
DB 181 AAGAGAAGAGAGTGCACAAACATCCCAACGCGGAAGCGCCCTCGTGGAACGCGAATCCCA 240
QY 3841 AGTTACCCACTAGCAAGGAGAGCCCACTCACTGCTCCGAGTGGGCAAGCTTTCAGAA 3900
DB 241 AGTTACCCACTAGCAAGGAGAGCCCACTCACTGCTCCGAGTGGGCAAGCTTTCAGAA 300
QY 3901 CCTACCAACAGCTGGTCTTGCACTCCAGGGTCCCAAGAGAGACCGGAGGCGCGGCGG 3960
DB 301 CCTACCAACAGCTGGTCTTGCACTCCAGGGTCCCAAGAGAGACCGGAGGCGCGGCGG 360
QY 3961 AGTCGCCACCATGTCTGTGGAACGAGGAGCAGCGGGGACGTGTCTCTGACCTCGCCG 4020
DB 361 AGTCGCCACCATGTCTGTGGAACGAGGAGCAGCGGGGACGTGTCTCTGACCTCGCCG 420
QY 4021 CCCTCTGGATGAATGAGGCGGTGGATCGAGGGGAGGTGTCTGGAAGACGATCTG 4080
DB 421 CCCTCTGGATGAATGAGGCGGTGGATCGAGGGGAGGTGTCTGGAAGACGATCTG 480
QY 4081 AGGATGGGCTTCCGGAAGGAATCCATCTGGGTAAAGCTGCCCTGTCTCCGTCGCGTGT 4140
DB 481 AGGATGGGCTTCCGGAAGGAATCCATCTGGGTAAAGCTGCCCTGTCTCCGTCGCGTGT 540
QY 4141 TCGCCTGTGTCTGTCTGCCGCTTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4200
DB 541 TCGCCTGTGTCTGTCTGCCGCTTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
QY 4201 TGGCCAGGAATGGGTTTGGAGAGCCAGAGTCAGTCCAGGCTCTTTTGGTATCACTCT 4260
DB 601 TGGCCAGGAATGGGTTTGGAGAGCCAGAGTCAGTCCAGGCTCTTTTGGTATCACTCT 660
QY 4261 GTGTAAGTCATTTAACTCTCAGGGCTTAAATTTTCTCATTTCTGTAATAACAGGGTTGA 4320
DB 661 GTGTAAGTCATTTAACTCTCAGGGCTTAAATTTTCTCATTTCTGTAATAACAGGGTTGA 720
QY 4321 GTTAAGAGTCTCCTTGTCTGAAAATATATATATATATATATATATATATATATATAT 4380
DB 721 GTTAAGAGTCTCCTTGTCTGAAAATATATATATATATATATATATATATATATATAT 780
QY 4381 TCACAAAACACATTTAAAAAATAAATCTGTGATCCAGGCCAAATGCACTGCTTCT 4440
DB 781 TCACAAAACACATTTAAAAAATAAATCTGTGATCCAGGCCAAATGCACTGCTTCT 840
QY 4441 TAACTGGGCGATTTTGTTCCTCAATCAGTATCTGGCAATCTCTGGAGGCAATTTGTGT 4500
DB 841 TAACTGGGCGATTTTGTTCCTCAATCAGTATCTGGCAATCTCTGGAGGCAATTTGTGT 900
QY 4501 CATACTGTGTGTGGGTGTGCTGTGGCATCCAGTGGGAGAGGCCAGGGACACTGCT 4560
DB 901 CATACTGTGTGTGGGTGTGCTGTGGCATCCAGTGGGAGAGGCCAGGGACACTGCT 960
QY 4561 CAGCATGTTACAGTGACAGGACAGCCCATCATCAAGAAATATCTGTCTCCAAATGTC 4620
DB 961 CAGCATGTTACAGTGACAGGACAGCCCATCATCAAGAAATATCTGTCTCCAAATGTC 1020
QY 4621 AATAGTTTTCAGCATTCAGAGACCTAGCTTCACTTAAAGTTTCTGGCGTTCCTGATCT 4680
DB 1021 AATAGTTTTCAGCATTCAGAGACCTAGCTTCACTTAAAGTTTCTGGCGTTCCTGATCT 1080
QY 4681 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAGGTACTGGGAG--TGATCAACTAGAGC 4737

DB 1081 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAGGTACTGGAGGTGATGATCAACTAGAGC 1140
QY 4738 CAGGAATATATTTGGGACGCGTTTGGTCTGTCTCAGAAACCTTGTCTTCTGTCTGGC 4797
DB 1141 CAGGAATATATTTGGGACGCGTTTGGTCTGTCTCAGAAACCTTGTCTTCTGTCTGGC 1200
QY 4798 AAGCTAGTATCCATTTATAGGTACCTCAGGAAACCAAAATGATTTGTCTATAAATAACAAG 4857
DB 1201 AAGCTAGTATCCATTTATAGGTACCTCAGGAAACCAAAATGATTTGTCTATAAATAACAAG 1260
QY 4858 AATGTGAGCACACTGAAGACATTTTAAAGAGCTCATTTGCTCAGCAGAAATTTTCAGTG 4917
DB 1261 AATGTGAGCACACTGAAGACATTTTAAAGAGCTCATTTGCTCAGCAGAAATTTTCAGTG 1320
QY 4918 TACTAGTGGCATTTATAGAAGAGAGAGGTGATCACTGAAGGCATGCTCATATAATATTC 4977
DB 1321 TACTAGTGGCATTTATAGAAGAGAGAGGTGATCACTGAAGGCATGCTCATATAATATTC 1380
QY 4978 TGAGCCCTGTGGGCGTTATCTAGGCAAAAGGATTCACCTGTGTTTGGAGTTGCGGCCCA 5037
DB 1381 TGAGCCCTGTGGGCGTTATCTAGGCAAAAGGATTCACCTGTGTTTGGAGTTGCGGCCCA 1440
QY 5038 TCCTCACTGTAGCCAGAGCTTCTCTATCAGAGTTTAGTATTTTGTGTAATAGAGGATC 5097
DB 1441 TCCTCACTGTAGCCAGAGCTTCTCTATCAGAGTTTAGTATTTTGTGTAATAGAGGATC 1500
QY 5098 TTGCTCTTAAAAACAGTTGAAAAGACCTGTAGGGCAGGCGTAATTTGACAGGGAATGA 5157
DB 1501 TTGCTCTTAAAAACAGTTGAAAAGACCTGTAGGGCAGGCGTAATTTGACAGGGAATGA 1560
QY 5158 TGGGAACATGAATCGGCTTTAGGGAAGCATCTCTCAAAAGTGCTCTTGGTTAAAAACAAGT 5217
DB 1561 TGGGAACATGAATCGGCTTTAGGGAAGCATCTCTCAAAAGTGCTCTTGGTTAAAAACAAGT 1620
QY 5218 GCTCTCTCTCTCAGTGTCACTTGAATTTGTGTGTTGAAATTTCTCGAAAACTCGGTGTAT 5277
DB 1621 GCTCTCTCTCTCAGTGTCACTTGAATTTGTGTGTTGAAATTTCTCGAAAACTCGGTGTAT 1680
QY 5278 GAGACCCAGATGAATTTGCCCAACAGATTTGATGAGCTCTTCCCTTCCAGCTCTTTCAG 5337
DB 1681 GAGACCCAGATGAATTTGCCCAACAGATTTGATGAGCTCTTCCCTTCCAGCTCTTTCAG 1740
QY 5338 CAGTGCAGTGTCTCTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 5397
DB 1741 CAGTGCAGTGTCTCTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
QY 5398 ATCTTTAGAAATGTTTGTAGTTTCTTGGGACACAGGAAACCCAGCACTTAGCATACTACA 5457
DB 1801 ATCTTTAGAAATGTTTGTAGTTTCTTGGGACACAGGAAACCCAGCACTTAGCATACTACA 1860
QY 5458 AATCTAATGCTTAATGGCATCATATAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAA 5517
DB 1861 AATCTAATGCTTAATGGCATCATATAAAGAGGCTTTAAACACAGACTCCAGTTAGCTAA 1920
QY 5518 GTGGTTTCTGTAGTGCCTGCTGTGTGAGGGGCTCTGTGAGATGCCCGCTTCCCTGA 5577
DB 1921 GTGGTTTCTGTAGTGCCTGCTGTGTGAGGGGCTCTGTGAGATGCCCGCTTCCCTGA 1980
QY 5578 AAGAAATGAAAGGCGAGTTACCGGTAGGTGTGTGGAATAACATGGGCTAGATCATCAGG 5637
DB 1981 AAGAAATGAAAGGCGAGTTACCGGTAGGTGTGTGGAATAACATGGGCTAGATCATCAGG 2040
QY 5638 CAGGACAGATGCTGCTGGCTGTGGGTGGGAGCACCCAGCTTGGCGTTGAGTTCTGTTCT 5697
DB 2041 CAGGACAGATGCTGCTGGCTGTGGGTGGGAGCACCCAGCTTGGCGTTGAGTTCTGTTCT 2100
QY 5698 ACCACTGCTGCTTTGTGACCAATTTATCAGTTGCTTAACTTTCTTTCTGCTACTATTTCC 5757
DB 2101 ACCACTGCTGCTTTGTGACCAATTTATGAGTTGCTTAACTTTCTTTCTGCTACTATTTCC 2160
QY 5758 CTGTTTGGAAAAATGTTTCAATGACCCCTGTTTCCACCTCCCAAGGACAATTTCAACAGC 5817

QY 7978 AGTTATTTAACTTTTGTGCTTTTGTCTTTCTTTGCTGTAAAGTAGAATAATACATAT 8037
DB AGTTATTTAACTTTTGTGCTTTTGTGCTTTTGTGCTGTAAAGTAGAATAATACATAT 4440
QY 8038 TTCCCTAGGCTCTTAGGAGATTAATAGTTAGAGTGTGCTGTAAATTTTCTATT 8097
DB TTCCCTAGGCTCTTAGGAGATTAATAGTTAGAGTGTGCTGTAAATTTTCTATT 4500
QY 8098 GAAGATAGGCATTTCAATAATTTCAATATTCATTTACAGTAAAGTAGAATAAGAACTGATGA 8157
DB GAAGATAGGCATTTCAATAATTTCAATATTCATTTACAGTAAAGTAGAATAAGAACTGATGA 4560
QY 8158 GAAATCCTATGTGATAGTAGATCGAAGAACAAAGAGAGAAAGACCTGTTTCTTAA 8217
DB GAAATCCTATGTGATAGTAGATCGAAGAACAAAGAGAGAAAGACCTGTTTCTTAA 4620
QY 8218 TAAATAGATATTTGATCTATTTTCAAGTCTTTTATACACTTCTTATATAAGTGCATTT 8277
DB TAAATAGATATTTGATCTATTTTCAAGTCTTTTATACACTTCTTATATAAGTGCATTT 4680
QY 8278 CTTGCTTAGGTCGAAACCAATGTAATTTTGTGAATATGCTGAGCCCAAGAAG 8337
DB CTTGCTTAGGTCGAAACCAATGTAATTTTGTGAATATGCTGAGCCCAAGAAG 4740
QY 8338 ACATCTCTGAGTATCACTTGGAGAGACATCAAGAGAAACAAACCGATGTTGCTGCT 8397
DB ACATCTCTGAGTATCACTTGGAGAGACATCAAGAGAAACAAACCGATGTTGCTGCT 4800
QY 8398 GAAGTCAAGAACGATGTTAAATACAGGACACTGAAGATGCACTATTAACCGCTGACAGT 8457
DB GAAGTCAAGAACGATGTTAAATACAGGACACTGAAGATGCACTATTAACCGCTGACAGT 4860
QY 8458 GCGCAACCAAAATTTGAAAGATTTTGTGAGTGCCAAAGATGTTACAGGCACTCCA 8517
DB GCGCAACCAAAATTTGAAAGATTTTGTGAGTGCCAAAGATGTTACAGGCACTCCA 4920
QY 8518 CTTGCAAAACGAGTTAAGAGATGCTTCTGTTTTTTCAGAAATGTTCTGGGCGAGCGTGT 8577
DB CTTGCAAAACGAGTTAAGAGATGCTTCTGTTTTTTCAGAAATGTTCTGGGCGAGCGTGT 4980
QY 8578 CTCTACAGGACACAAAGATCTAGGATTTCCATATAAATGACAGTATGACAGTGT 8637
DB CTCTACAGGACACAAAGATCTAGGATTTCCATATAAATGACAGTATGACAGTGT 5040
QY 8638 GATAAGTGAATAAAACCTTACCTGCTTACCTGCACTGTTAAAGAGATCAGCA 8697
DB GATAAGTGAATAAAACCTTACCTGCTTACCTGCACTGTTAAAGAGATCAGCA 5100
QY 8698 GTTGAAACTCAGGCAAAATCACTCTGTAGAACCAAGCGAGATGTTACTCTCTCCG 8757
DB GTTGAAACTCAGGCAAAATCACTCTGTAGAACCAAGCGAGATGTTACTCTCTCCG 5160
QY 8758 GATGCGAGTACACCCATTAACCTTGAAGTTAGCCCAAGAGAACCAAGCGAGACCGCA 8817
DB GATGCGAGTACACCCATTAACCTTGAAGTTAGCCCAAGAGAACCAAGCGAGACCGCA 5220
QY 8818 GCTGACTGAGATACAGGCAAGTGTGATTTGACGAAACCTTTAAATTTTCCGTT 8877
DB GCTGACTGAGATACAGGCAAGTGTGATTTGACGAAACCTTTAAATTTTCCGTT 5280
QY 8878 GGGGCTCTTCAATTTGCCCGCAATTTCTTTGAGTAAAGTTTGATTTCAAGTATCACC 8937
DB GGGGCTCTTCAATTTGCCCGCAATTTCTTTGAGTAAAGTTTGATTTCAAGTATCACC 5340
QY 8938 TGTCATTTTGTACCTTCAAGCATTTTATCCAGAGTTTATGATGACCAAGAGACTG 8997
DB TGTCATTTTGTACCTTCAAGCATTTTATCCAGAGTTTATGATGACCAAGAGACTG 5400
QY 8998 GAGCATAAATCAATCTGACCTTCATAAAACCTGCGAAACAGTCTTGTGCTTAAAGT 9057
DB GAGCATAAATCAATCTGACCTTCATAAAACCTGCGAAACAGTCTTGTGCTTAAAGT 5460
QY 9058 CGAGCTACCGGATGCCCGCAGCTGTGCTGGGAAAGATGTGCTTCCCTCCCTCTAGTTTC 9117

DB CGAGCTACCGGATGCCCGCAGCTGTGCTGGAAAGATGTGCTTCTCTAGTTTC 5520
QY 9118 TGTAAACCCAGCCCAAGTCTGCTTTCCCGGCGAGTCCAAATCCCTGCGCATCTGCGAAG 9177
DB TGTAAACCCAGCCCAAGTCTGCTTTCCCGGCGAGTCCAAATCCCTGCGCATCTGCGAAG 5580
QY 9178 GCGAAGCAGAGCCCTCTCTGGCCAGGCAAGCCCTCTGACTTTCAGGGATAGACTCTAGC 9237
DB GCGAAGCAGAGCCCTCTCTGGCCAGGCAAGCCCTCTGACTTTCAGGGATAGACTCTAGC 5640
QY 9238 ACTTTAGCCCCAGTAACCTGAACTCCACAGACCAAGAGAAATGTGGGGGTCCTAAGG 9297
DB ACTTTAGCCCCAGTAACCTGAACTCCACAGACCAAGAGAAATGTGGGGGTCCTAAGG 5700
QY 9298 GCGCCACACAGGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCCCTGCAACG 9357
DB GCGCCACACAGGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCCCTGCAACG 5760
QY 9358 GATAAGACAAAAGACCCGAGACAAATTTGAAACCTTCCAGTAGTCTTCTCAGCCCC 9417
DB GATAAGACAAAAGACCCGAGACAAATTTGAAACCTTCCAGTAGTCTTCTCAGCCCC 5820
QY 9418 ACCCTCGGACAGTAACATCAATGTTTCCATCGACTACCCGCCCAAGAACGACGAGCCG 9477
DB ACCCTCGGACAGTAACATCAATGTTTCCATCGACTACCCGCCCAAGAACGACGAGCCG 5880
QY 9478 TGGGCACTCTCGGGAAGAGACTATTTCTGTAATCGAGTGCAGCAATCTGCGAGCAAA 9537
DB TGGGCACTCTCGGGAAGAGACTATTTCTGTAATCGAGTGCAGCAATCTGCGAGCAAA 5940
QY 9538 TTTTGTGAGCCCTTCCAAAAGACTGAAAGTCCAGCGTGGTCCCTTCAGCTTGACCA 9597
DB TTTTGTGAGCCCTTCCAAAAGACTGAAAGTCCAGCGTGGTCCCTTCAGCTTGACCA 6000
QY 9598 CCCGGGGCCAAATTACAGAAAGAGCTATGACTTTCCTCAAGTACCATATGCTCAGAGCATC 9657
DB CCCGGGGCCAAATTACAGAAAGAGCTATGACTTTCCTCAAGTACCATATGCTCAGAGCATC 6060
QY 9658 ACATCACTGTTACCGAGAGCTGTGTATTCGTCGAGCGCTGCTCCCAAAACCAAGG 9717
DB ACATCACTGTTACCGAGAGCTGTGTATTCGTCGAGCGCTGCTCCCAAAACCAAGG 6120
QY 9718 TTTCTGAGCTCCAGCGAGTCCGATTTCTCCAAATGTCTGACTGTTTCAAGAGCCCTATGCT 9777
DB TTTCTGAGCTCCAGCGAGTCCGATTTCTCCAAATGTCTGACTGTTTCAAGAGCCCTATGCT 6180
QY 9778 GGCTCCGGGCCACTTTTACACTTTGTGTGCTGCTGTGCTAGTCCAGCATCCAGCTCGAGTTA 9837
DB GGCTCCGGGCCACTTTTACACTTTGTGTGCTGCTGTGCTAGTCCAGCATCCAGCTCGAGTTA 6240
QY 9838 GAAGGTATTGCATGAGGGGCGTGTGTTTAAATGCGCTGCTACAGTATTAATAGCTAAT 9897
DB GAAGGTATTGCATGAGGGGCGTGTGTTTAAATGCGCTGCTACAGTATTAATAGCTAAT 6300
QY 9898 CCAGGCACTTCAGTGGAGTGGTACCACTCCCAAGGGTGGGGGTAGGAGCCCAAGT 9957
DB CCAGGCACTTCAGTGGAGTGGTACCACTCCCAAGGGTGGGGGTAGGAGCCCAAGT 6360
QY 9958 TCTTGGGGGTCCAGAGAGAAAGCAATTTCTAGATAGCGGAGTGG 10000
DB TCTTGGGGGTCCAGAGAGAAAGCAATTTCTAGATAGCGGAGTGG 6403

RESULT 4

ACC50334

ID ACC50334 standard; cdna; 5632 BP.

XX ACC50334;

AC ACC50334;

XX 12-JUN-2003 (first entry)

XX

DE Breast cancer associated cDNA sequence SEQ ID NO:505.

XX Human; breast cancer; cytostatic; gene therapy; gene; ss.
 XX Homo sapiens.
 XX WO2003004989-A2.
 XX 16-JAN-2003.
 XX 21-JUN-2002; 2002WO-US019669.
 XX 21-JUN-2001; 2001US-0299887P.
 XX 27-JUN-2001; 2001US-0301572P.
 XX 18-JUL-2001; 2001US-0306501P.
 XX 25-SEP-2001; 2001US-0325002P.
 XX 05-MAR-2002; 2002US-0362585P.
 XX 14-MAY-2002; 2002US-0380391P.
 XX (MILL-) MILLENIUM PHARM INC.
 XX Lillie J, Gannavarapu M, Glatt K, Hoersh S, Kamatkar S;
 PI Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;
 PI Bast RC, Horobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;
 XX
 DR WPI: 2003-210381/20.
 DR P-PSDB; ABR47632.
 XX
 PT Breast cancer diagnosis or treatment by comparing the level of expression
 PT of a marker in a patient sample with that in the control non-breast
 PT cancer sample.
 XX
 PS Claim 1; SEQ ID NO 505; 128pp; English.
 XX
 CC The present invention describes a method for assessing whether a patient
 CC is afflicted with breast cancer. The method comprises comparing the level
 CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
 CC ABR47386 to ABR47632) in a patient sample and the normal level of
 CC expression of the marker in a control non-breast cancer sample, where a
 CC significant increase in the level of expression of the marker in the
 CC patient sample and the normal level is an indication that the patient is
 CC afflicted with breast cancer. The breast cancer associated sequences from
 CC the present invention have cytostatic activities and can be used in gene
 CC therapy. The method is useful for diagnosing and treating breast cancer.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 5632 BP; 1637 A; 1166 C; 1258 G; 1571 T; 0 U; 0 Other;
 Query Match 16.4%; Score 1639.4; DB 8; Length 5632;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1640; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2474 GACAAAGAGAACTAATGCTTTGCTGCTGATTATTTGAATCGAGGCATTGGGAACCCCTG 2533
 DB 1 GACAAAGAGAACTAATGCTTTGCTGCTGATTATTTGAATCGAGGCATTGGGAACCCCTG 60
 QY 2534 TATGCTTTGTTGTGGAAGAACCAAGTGCACCATCAGCTTCTTAAAGTTTGAAG 2593
 DB 61 TATGCTTTGTTGTGGAAGAACCAAGTGCACCATCAGCTTCTTAAAGTTTGAAG 120
 QY 2594 AAGTTAGAGGACTATACACTTTCTTTGAACTTTTATAATAATTTGCTCTGTTT 2653
 DB 121 AAGTTAGAGGACTATACACTTTCTTTGAACTTTTATAATAATTTGCTCTGTTT 180
 QY 2654 GGAACCCAGGGCTTTAGAGGGGTGAGTGCACCAAGTCTTACAAGTGGCTTTATTCACATC 2713
 DB 181 GGAACCCAGGGCTTTAGAGGGGTGAGTGCACCAAGTCTTACAAGTGGCTTTATTCACATC 240
 QY 2714 CAGAAATGCCCCAAGCGAACTTTTGAAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 2773
 DB 241 CAGAAATGCCCCAAGCGAACTTTTGAAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300

QY 2774 AACTCAATCCCTCTTAATGTATCATGGATGGCCAGAGTGTGGCAGCTCTCTTGGCAG 2833
 DB 301 AACTCAATCCCTCTTAATGTATCATGGATGGCCAGAGTGTGGCAGCTCTCTTGGCAG 360
 QY 2834 TCCGATGGAGATGGAGGATGCTTGTCAATGAAAGGGACCGCTGTTGTTTCCATTCGGAGC 2893
 DB 361 TCCGATGGAGATGGAGGATGCTTGTCAATGAAAGGGACCGCTGTTGTTTCCATTCGGAGC 420
 QY 2894 TACACAAGAAAAAATGTATCATCCAAATCGAGGGGTATATGCCCTTGGATGTGATGTTCTG 2953
 DB 421 TACACAAGAAAAAATGTATCATCCAAATCGAGGGGTATATGCCCTTGGATGTGATGTTCTG 480
 QY 2954 CAGCCAGACCTTTCACACATTCAGAGACCTTAATAAACAATGCTTTAAATGCAACACCGGCC 3013
 DB 481 CAGCCAGACCTTTCACACATTCAGAGACCTTAATAAACAATGCTTTAAATGCAACACCGGCC 540
 QY 3014 TACCTCTGTGAAACCCAGCAGTCTTTCGGGTGTGAAGCAGAGTATCTCAGTCCGCTTTGATAA 3073
 DB 541 TACCTCTGTGAAACCCAGCAGTCTTTCGGGTGTGAAGCAGAGTATCTCAGTCCGCTTTGATAA 600
 QY 3074 AAGTCAAGTCGACACAGAACCTCCCAAGAAAAAAGAAATTCAGAGAAAAATGAATTTAGCTG 3133
 DB 601 AAGTCAAGTCGACACAGAACCTCCCAAGAAAAAAGAAATTCAGAGAAAAATGAATTTAGCTG 660
 QY 3134 TGAGGTATGTGGGCGACATTTAGAGTGCCTTTTATGTTGAGATCCCATGAGAAACACA 3193
 DB 661 TGAGGTATGTGGGCGACATTTAGAGTGCCTTTTATGTTGAGATCCCATGAGAAACACA 720
 QY 3194 CAAAGATTCTTTTCACTTACGGGTGTAAACATGTGCGGAAAGAAATTCAGAGGACCTTGGTT 3253
 DB 721 CAAAGATTCTTTTCACTTACGGGTGTAAACATGTGCGGAAAGAAATTCAGAGGACCTTGGTT 780
 QY 3254 TCTTAAAAATCACATGCGGACACATAATGCGGAAATTCGGGGGCGACAGAAACCTGACGCA 3313
 DB 781 TCTTAAAAATCACATGCGGACACATAATGCGGAAATTCGGGGGCGACAGAAACCTGACGCA 840
 QY 3314 AGGCTTGAGAGTAGTCCAGCAACCATCAACGAGTCTGTCAGAGTGCACGCGGCGGAG 3373
 DB 841 AGGCTTGAGAGTAGTCCAGCAACCATCAACGAGTCTGTCAGAGTGCACGCGGCGGAG 900
 QY 3374 CATCTCTCTCTCTTCAAAAATCTGATGCTTGTGCTTCTTATTTTCAAAATAAAGAAAG 3433
 DB 901 CATCTCTCTCTCTTCAAAAATCTGATGCTTGTGCTTCTTATTTTCAAAATAAAGAAAG 960
 QY 3434 TCTAATGAGCAACCGCAAGGTGCACACAAAAAAACTGCTTTCGGTACCGAGCAGCGCA 3493
 DB 961 TCTAATGAGCAACCGCAAGGTGCACACCAAAAAACTGCTTTCGGTACCGAGCAGCGCA 1020
 QY 3494 GACAGACTCTCCACAAGGAGGAATCCGCTCTCGAGGGAGGACTTCTCGAGTGTTCAA 3553
 DB 1021 GACAGACTCTCCACAAGGAGGAATCCGCTCTCGAGGGAGGACTTCTCGAGTGTTCAA 1080
 QY 3554 CTTGAGACCAAAATCTCACCTTGAAACCGGGGAAGAGCCTGTGCATGTCATCCCTCAGCT 3613
 DB 1081 CTTGAGACCAAAATCTCACCTTGAAACCGGGGAAGAGCCTGTGCATGTCATCCCTCAGCT 1140
 QY 3614 CGATCCGTTTACCACTTCCAGGCTTGGCAGTGTGCTGCTACCAAGAAAGTTGCCATTG 3673
 DB 1141 CGATCCGTTTACCACTTCCAGGCTTGGCAGTGTGCTGCTACCAAGAAAGTTGCCATTG 1200
 QY 3674 CCAAGAAGTGAAGGAATCGGGCAAGAGGGAGCAGCAGCAACACGATTCGAGTTCGCA 3733
 DB 1201 CCAAGAAGTGAAGGAATCGGGCAAGAGGGAGCAGCAGCAACACGATTCGAGTTCGCA 1260
 QY 3734 GAAGGAGCTTGGAGAAAAAATAAGGCGAGTTGTGCGAGCCTCTCGCAAGAGAAAGAGAA 3793
 DB 1261 GAAGGAGCTTGGAGAAAAAATAAGGCGAGTTGTGCGAGCCTCTCGCAAGAGAAAGAGAA 1320
 QY 3794 GTGCAAAACATCTCCACGCGGAAGCGCTTCCGTGAGCGGGATCCCAAGTTACCCAGTAG 3853
 DB 1321 GTGCAAAACATCTCCACGCGGAAGCGCTTCCGTGAGCGGGATCCCAAGTTACCCAGTAG 1380
 QY 3854 CAAGGAGAAAGCCCACTCACTGCTCGAGTGGCGCAAGAGCTTTTTCAGAACCTACCAACAGCT 3913

Db	1381	CAAGGAGAGCCCACTACTGCTCGAGTCGGGCAAGCTTTGAGAACCTTACACACGACT	1440
Qy	3914	GGTCTTGCATCTCAGAGGTCCACAAGAAGACCGAGAGGGCGGCGGAGTGCGCCACCAT	3973
Db	1441	GGTCTTGCATCTCAGAGGTCCACAAGAAGACCGAGAGGGCGGCGGAGTGCGCCACCAT	1500
Qy	3974	GTCTGTGGAACGGGAGGACGCGGGGACGTTGTTCTCTGACCTCGCCGCCCTCTGGATGA	4033
Db	1501	GTCTGTGGAACGGGAGGACGCGGGGACGTTGTTCTCTGACCTCGCCGCCCTCTGGATGA	1560
Qy	4034	AAATGGAGCCGTGGATCGAGGGGAGGTGTTCTGAGAGACGGATCTGAGGATGGGCTTCC	4093
Db	1561	AAATGGAGCCGTGGATCGAGGGGAGGTGTTCTGAGAGACGGATCTGAGGATGGGCTTCC	1620
Qy	4094	CGAAGGAATCCATCTGGGTAA	4114
Db	1621	CGAAGGAATCCATCTGGATAA	1641

RESULT 5

ADD14635

ID ADD14635 standard; cDNA; 5632 BP.

XX

AC ADD14635;

XX
DT 01-JAN-2004 (first entry)

XX
DE Human src biomarker polynucleotide SEQ ID NO:29.

XX predictor set; protein tyrosine kinase activity modulator;
KW protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
KW gene therapy; drug sensitivity; genetic profile; cancer; human; gene; ss.

XX
OS Homo sapiens.

XX PN WO2003062395-A2.

31-JUL-2003

XX PF 17-JAN-2003; 2003WO-US001981.

18-JAN-2002: 2002US-0350061P.

PA (BRIM) BRISTOL-MYERS SQUIBB CO.

XX
PI Huang F, Fairchild CR, Lee FY, Shaw P;

XX
DR WPI; 2003-636735/60.
DR P-PSDB; ADD14040.

XX New polynucleotides and polypeptides for predicting the activity of
PT compounds that interact with protein tyrosine kinases and/or protein
PT tyrosine kinase pathways.

XX
PS Claim 2; SEQ ID NO 29; 139pp; English.

The present invention describes a predictor set comprising a plurality of polynucleotides or polypeptides whose expression pattern is predictive of the response of cells to treatment with a compound that modulates protein tyrosine kinase activity or members of the protein tyrosine kinase pathway. Also described: (1) predicting whether a compound is capable of modulating the activity of cells, comprising obtaining a sample of cells, determining whether the cells express a plurality of markers, and correlating the expression of the markers to the compound's ability to modulate the activity of the cells; (2) a plurality of cell lines for identifying polynucleotides and polypeptides whose expression levels correlate with compound sensitivity or resistance of cells associated with a disease state; and (3) identifying polynucleotides and polypeptides that predict compound sensitivity or resistance of cells associated with a disease state, comprising subjecting the plurality of cell lines to one or more compounds, analysing the expression pattern of a microarray of polynucleotides or polypeptides, and selecting

[illegible]

Db	1201	CCAAAGAGTGAAGGAATCGGGCAAGAGGAGCACCACAACGACGATTCGAGTTCCGA	1260
Qy	3734	GAAGGAGCTTGGAGAPACAATAAGGACGCTGTGCAGGCCTCTCGCAAGAGAAAGAA	3793
Db	1261	GAAGGAGCTTGGAGAAAACAATAAGGACGCTGTGCAGGCCTCTCGCAAGAGAAAGAA	1320
Qy	3794	GTGCAAAACACTCCACAGGCAAGCGCCCTCCGTGGAGCGGATCCCAAGTTACCAGTAG	3853
Db	1321	GTGCAAAACACTCCACAGGCAAGCGCCCTCCGTGGAGCGGATCCCAAGTTACCAGTAG	1380
Qy	3854	CAAGGAGAAGCCCACTCACTGCTCCGAGTCGGCAAAAGCTTTCAGAACCTTACCACAGCT	3913
Db	1381	CAAGGAGAAGCCCACTCACTGCTCCGAGTCGGCAAAAGCTTTCAGAACCTTACCACAGCT	1440
Qy	3914	GGTCTTGCACTCCAAGGTTCCAAGAAGGACCGAGGGCGCGGCGAGTCGCCCAACCAT	3973
Db	1441	GGTCTTGCACTCCAAGGTTCCAAGAAGGACCGAGGGCGCGGCGAGTCGCCCAACCAT	1500
Qy	3974	GTCTGTGACGGAGGAGCAGCCGGGACGTGTTCTCTGACCTCGCCGCCCTCTGGATGA	4033
Db	1501	GTCTGTGACGGAGGAGCAGCCGGGACGTGTTCTCTGACCTCGCCGCCCTCTGGATGA	1560
Qy	4034	AAATGAGCCGTGGATCGAGGGGAAGTGTTCTTGAGACGATCTTGAGATGGGCTTCC	4093
Db	1561	AAATGAGCCGTGGATCGAGGGGAAGTGTTCTTGAGACGATCTTGAGATGGGCTTCC	1620
Qy	4094	CGAAGGAATCCATCTGGGTHA	4114
Db	1621	CGAAGGAATCCATCTGGATAA	1641

RESULT 7

AD	ADR65858	standard; DNA; S632 BP.
ID	ADR65858	
XX	AC	
XX	AC	
XX	ADR65858;	
XX	AC	
XX	02-DEC-2004	(first entry)
XX	DT	
XX	DE	Human prostatic carcinoma derived DNA SEQ ID 54 #1.
XX	XX	
XX	KW	human; cytostatic; diagnosis; prostatic cancer;
XX	KW	differential expression analysis; ds.
XX	XX	
XX	OS	Homo sapiens.
XX	XX	
XX	FN	WO2004076614-A2.
XX	XX	
XX	PD	10-SEP-2004.
XX	XX	
XX	PF	22-FEB-2004; 2004WO-DE000433.
XX	XX	
XX	PR	27-FEB-2003; 2003DE-01009985.
XX	PR	14-MAY-2003; 2003DE-01022134.
XX	XX	
XX	PA	(HINZ/) HINZMANN B.
XX	PA	(DAHL/) DAHL E.
XX	PA	(ROSE/) ROSENTHAL A.
XX	PA	(HERM/) HERMANN K.
XX	PA	(PILA/) PILARSKY C.
XX	XX	
XX	PI	Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pil
XX	PI	Schmitt A, Beckmann G, Bruemendorf T, Kinnemann
XX	PI	Xinzhong L, Staub E;
XX	XX	
XX	DR	WPI; 2004-653386/63.
XX	XX	
XX	PT	New nucleic acids, and encoded proteins, from prosta
XX	PT	useful for diagnosis, treatment and in screening fo
XX	PT	agents.
XX	XX	
XX	PS	Claim 1; Page 253-254; 1607pp; German.
XX	XX	

CC This invention describes novel cytosolic polynucleotide and polypeptide
CC sequences which can be used in a method for diagnosing prostatic cancer
CC or the risk of developing prostatic cancer. Diagnosis is based on
CC determining over transcription or over expression of the sequences in
CC prostatic tissue. Screening for inhibitors of the sequences or detection
CC substances involves a binding assay, any compounds that bind are
CC selected, optionally after deconvolution of mixtures. Detection of a
CC predetermined minimum level of the reporter indicates the presence of a
CC tumour cells. Inhibitors can be chosen from antisense oligonucleotides,
CC short-interfering RNA or ribozymes; an organic molecule of molecular
CC weight below 5000, preferably 300, that binds to the polypeptide; an
CC aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the
CC polypeptide, preferably humanised or human; an anti-idiotypic, non-human
CC (monoclonal) antibody directed against Ab or any of the above derivatised
CC with a reporter group, cell toxin, immunostimulatory molecules and/or
CC radioisotope. The polynucleotides are identified in human prostatic
CC cancer by differential expression analysis, using DNA microarrays,
CC between normal and tumorous tissues, with (over)expression being detected
CC by quantitative PCR. Analysis of prostatic cancer samples showed that
CC CD24 was upregulated in many of them. Sections of tissue, isolated from
CC prostatic cancer patients, or subjects at risk, were incubated
CC sequentially with anti-human CD4 murine monoclonal antibodies;
CC biotinylated second antibody; streptavidin-conjugated horseradish
CC peroxidase and then diaminobenzidine as colour former (brown). The
CC samples were counterstained with hemalum (blue). Malignant cells stained
CC strongly but non-malignant cells only weakly. In 15 of 63 samples of
CC adenocarcinoma, membrane and cytoplasmic staining was very strong, and
CC lymph node metastases were also stained. ADR65805-ADR6954 represent the
CC polynucleotide and polypeptide sequences used in the method of the
CC invention.

XX
SQ Sequence 5632 BP; 1637 A; 1166 C; 1258 G; 1571 T; 0 U; 0 Other;

Query Match 16.4%; Score 1639.4; DB 13; Length 5632;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1640; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	2474	GACAAAGAGAACTAATGCTTTGCTGATTCATATTTGAATCGAGGCATTTGGGAAACCTTG	2533
DB	1	GACAAAGAGAACTAATGCTTTGCTGATTCATATTTGAATCGAGGCATTTGGGAAACCTTG	60
QY	2534	TATGCTCTTTGTGGAAGAACCACTGACACATCAGCTTCTTAAAGTTTCTGAAG	2593
DB	61	TATGCTCTTTGTGGAAGAACCACTGACACATCAGCTTCTTAAAGTTTCTGAAG	120
QY	2594	AAAGTTAGAGGACTATACACTTTCTTTTGAACCTTTTATATAAATATTTGCTCTGTTTTT	2653
DB	121	AAAGTTAGAGGACTATACACTTTCTTTTGAACCTTTTATATAAATATTTGCTCTGTTTTT	180
QY	2654	GGAAACCCAGGCTGTAGAGGGGTGAGTGAACAGCTTCAAGTGGCCCTTATCCAACTC	2713
DB	181	GGAAACCCAGGCTGTAGAGGGGTGAGTGAACAGCTTCAAGTGGCCCTTATCCAACTC	240
QY	2714	CAGAAATTCGCCAACGAACTTTGAGATTATATGCAATCGAAAGTGACAGAAACATGCC	2773
DB	241	CAGAAATTCGCCAACGAACTTTGAGATTATATGCAATCGAAAGTGACAGAAACATGCC	300
QY	2774	AACTCAATCCCTCTTAATGTATCATGGATGGCCAGAAAGTGATGGCAGCTCTCTTGGCAG	2833
DB	301	AACTCAATCCCTCTTAATGTATCATGGATGGCCAGAAAGTGATGGCAGCTCTCTTGGCAG	360
QY	2834	TCCGATGGAGATGGAGATGCCCTTGTCAATGAAGAGGACCGCTGTTCATTCGGAGC	2893
DB	361	TCCGATGGAGATGGAGATGCCCTTGTCAATGAAGAGGACCGCTGTTCATTCGGAGC	420
QY	2894	TACAAAGAAAAATGTTCATCCAAATCGAGGGGTATATGCCCTTGGATTGATGTTCTG	2953
DB	421	TACAAAGAAAAATGTTCATCCAAATCGAGGGGTATATGCCCTTGGATTGATGTTCTG	480
QY	2954	CAGCCAGACCTTCACACTTCAGAGAACCTTAAATAACATGCTTAAATGCAACACCGGCC	3013
DB	481	CAGCCAGACCTTCACACTTCAGAGAACCTTAAATAACATGCTTAAATGCAACACCGGCC	540

QY	3014	TACCTCTGTGAAACCCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA	3073
DB	541	TACCTCTGTGAAACCCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA	600
QY	3074	AAAGTCAAGTGGGAAACAGAAACCTCCCAAGGAAAAAATGCAAGGAAAAATGAATTTAGCTG	3133
DB	601	AAAGTCAAGTGGGAAACAGAAACCTCCCAAGGAAAAAATGCAAGGAAAAATGAATTTAGCTG	660
QY	3134	TGAGTATGTGGGACAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA	3193
DB	661	TGAGTATGTGGGACAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACACA	720
QY	3194	CAAGATTTCTTTCACTTACGGGTGTAAATGTGCGGAAAGAGATTCAAGGAGCTTTGGTT	3253
DB	721	CAAGATTTCTTTCACTTACGGGTGTAAATGTGCGGAAAGAGATTCAAGGAGCTTTGGTT	780
QY	3254	TCTTAAAAATCACATGCGGACACATAATGGCAAAATCGGGGGCCAGAAACAACTCAGCA	3313
DB	781	TCTTAAAAATCACATGCGGACACATAATGGCAAAATCGGGGGCCAGAAACAACTCAGCA	840
QY	3314	AGGCTTGAGAGTAGTCCAGCAACGATCAACGAGTCTGTCAGGTGCGACGCGCGGAGAG	3373
DB	841	AGGCTTGAGAGTAGTCCAGCAACGATCAACGAGTCTGTCAGGTGCGACGCGCGGAGAG	900
QY	3374	CATCTCTCTCTCTTACAAAATCTGATGTTGTGCTTCTCTTATTTCCAAAATAAAGAAAG	3433
DB	901	CATCTCTCTCTCTTACAAAATCTGATGTTGTGCTTCTCTTATTTCCAAAATAAAGAAAG	960
QY	3434	TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTTCGGTACCAAGCGCGCA	3493
DB	961	TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTTCGGTACCAAGCGCGCA	1020
QY	3494	GACAGACTCTCCAAGGAGGAATGCCGTCTCGAGGGAGGACTTCTGCGAGTTGTTCAA	3553
DB	1021	GACAGACTCTCCAAGGAGGAATGCCGTCTCGAGGGAGGACTTCTGCGAGTTGTTCAA	1080
QY	3554	CTTGAGACCAAAATCTCACCCCTGAAAACGGGGAAGAGCTGTGCAGATGCATCCCTCAGCT	3613
DB	1081	CTTGAGACCAAAATCTCACCCCTGAAAACGGGGAAGAGCTGTGCAGATGCATCCCTCAGCT	1140
QY	3614	CGATCCGTTCCACCTTTCCAGGCTTGGCAGCTGTGCTACCAAGAAAAAGTTGCCATTG	3673
DB	1141	CGATCCGTTCCACCTTTCCAGGCTTGGCAGCTGTGCTACCAAGAAAAAGTTGCCATTG	1200
QY	3674	CCAAGAGTGAAGAAATCGGGCAGAGGAGGACCGACACACAGATTCAGAGTTCCGA	3733
DB	1201	CCAAGAGTGAAGAAATCGGGCAGAGGAGGACCGACACACAGATTCAGAGTTCCGA	1260
QY	3734	GAGGAGCTTGGAGAAAACAAATAAGGGCAGTTGTGCGAGGCTCTCGCAAGAGAAAGAGAA	3793
DB	1261	GAGGAGCTTGGAGAAAACAAATAAGGGCAGTTGTGCGAGGCTCTCGCAAGAGAAAGAGAA	1320
QY	3794	GTGCAAAACATCTCCAAGGAGAGCCCTCTCGTGGACGCGGATCCCAAGTTACCCAGTAG	3853
DB	1321	GTGCAAAACATCTCCAAGGAGAGCCCTCTCGTGGACGCGGATCCCAAGTTACCCAGTAG	1380
QY	3854	CAAGAGAGCCCACTCACTGCTCGAGTCGGGCAAGAGTTTCAGAACTACCAAGCT	3913
DB	1381	CAAGAGAGCCCACTCACTGCTCGAGTCGGGCAAGAGTTTCAGAACTACCAAGCT	1440
QY	3914	GGTCTTGCATCTCAGGGTCCCAAGAGGACCGGAGGGCCGGGGAGTCCGCCACCAT	3973
DB	1441	GGTCTTGCATCTCAGGGTCCCAAGAGGACCGGAGGGCCGGGGAGTCCGCCACCAT	1500
QY	3974	GTCTGTGAGCGGAGGACCGGGGACGTTGTTCTCTGACCTCGCGCGCCCTCTGATGA	4033
DB	1501	GTCTGTGAGCGGAGGACCGGGGACGTTGTTCTCTGACCTCGCGCGCCCTCTGATGA	1560
QY	4034	AAATCGAGCCGTGATTCGAGGGGAGGTTGTTCTGAGACGATCTGAGGATGGGCTCC	4093
DB	1561	AAATCGAGCCGTGATTCGAGGGGAGGTTGTTCTGAGACGATCTGAGGATGGGCTCC	1620
QY	4094	CGAAGGAATCCATCTCGGGTAA	4114

Db 1621 CGAAGGAATCATCTGGATAA 1641
|||||

RESULT 8
ADP07267
ID ADP07267 standard; DNA; 5653 BP.
XX

ADP07267;

29-JUL-2004 (first entry)

Human ZNF217 DNA.

de; proliferative disease; breast; methylation; CpG; bisulfite; human.

Homo sapiens.

DE10255104-Al.

11-MAR-2004.

26-NOV-2002; 2002DE-01055104.

27-AUG-2002; 2002DE-01039313.

(EPiG-) EPIGENOMICS AG.

Maier S;

WPI; 2004-284340/27.

GENBANK; NM_006526.

Analyzing proliferative diseases of breast cells, useful e.g. for diagnosis, prognosis and treatment of breast cancer, by determining methylation status of specific genes.

Claim 1; Page; 22pp; German.

This invention describes a novel method of analysing proliferative diseases of breast cells by determining the methylation status of certain genes. The invention also describes nucleic acids or their complements, oligomers, especially oligonucleotides or peptide nucleic acid oligomers, that hybridise to, or are identical with, any of the nucleic acids, the preparation of an oligomer array for analysing proliferative diseases of breast cells that are associated with the methylation status of CpG dinucleotides of the genes by bonding at least one oligomer to a solid phase and a kit comprising a bisulfite reagent and the oligomer. The nucleic acids are genomic sequences (5'- and/or regulatory and/or CpG-rich regions). The base sequence of the oligomer includes at least one CpG island, especially with C in the middle third of the sequence. The process involves a genomic DNA sample treated chemically, specifically with a bisulfite reagent, to convert non-methylated C to uracil or some other base having base-pairing properties different from those of C. Fragments of the treated DNA are then amplified, using primers and a polymerase and the methylation status of the genomic CpG dinucleotides is determined by analysis of the amplicons, particularly by hybridisation to the oligomer, optionally with extension of the hybridised oligomer by at least one base, or detection is by sequencing. The amplification may use methylation-specific primers. Alternatively, genomic DNA is extracted from a sample and digested with methylation-specific restriction enzymes, then the digestion fragments detected, optionally after amplification. In either method, more than 10 fragments of 100-200 bp are amplified in a single reaction vessel, using a heat-resistant DNA polymerase in PCR. The amplicons carry detectable markers, e.g. fluorophores, radioisotopes and/or releasable fragments of known mass that can be detected by mass spectrometry. The method is used for characterisation, classification, differentiation, staging, phase-estimation, diagnosis and/or therapy of proliferative diseases of breast cells. The method provides very specific classification of proliferative diseases, allowing better treatment. It can both characterise methylation status and detect single-nucleotide polymorphisms. This sequence represents human gene used to illustrate the method of the invention. NOTE: This sequence does not appear in the

CC printed specification but has been retrieved from Genbank.

XX
SQ Sequence 5653 BP; 1656 A; 1167 C; 1258 G; 1572 T; 0 U; 0 Other;

Query Match 16.4%; Score 1639.4; DB 12; Length 5653;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1640; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2474 GACAAAGAGAACTAATGCTTTGTGCTGATTCAATATTGAATCGAGGCAATGGGAACCCCTG 2533
|||
DB 1 GACAAAGAGAACTAATGCTTTGTGCTGATTCAATATTGAATCGAGGCAATGGGAACCCCTG 60
QY 2534 TATGCTTGTGTGAAAGAACCACTGACACCATCACTGAGCTTCTTAAAGTTTGAAG 2593
|||
DB 61 TATGCTTGTGTGAAAGAACCACTGACACCATCACTGAGCTTCTTAAAGTTTGAAG 120
QY 2594 AGTTAGAGGACTATACACTTTCTTTTGAACCTTTTATATAAATATTGCTCTGCTTTT 2653
|||
DB 121 AGTTAGAGGACTATACACTTTCTTTTGAACCTTTTATATAAATATTGCTCTGCTTTT 180
QY 2654 GGAAACCCAGGGCTGTTAGAGGGGTGAGTGACAAAGTCTTACAAGTGGCTTTATTTCCAACTC 2713
|||
DB 181 GGAAACCCAGGGCTGTTAGAGGGGTGAGTGACAAAGTCTTACAAGTGGCTTTATTTCCAACTC 240
QY 2714 CAGAAATGCCAAACCGAACTTTTGAGATTATATGCAATCGAAAGTGCAGGAAACATGCC 2773
|||
DB 241 CAGAAATGCCAAACCGAACTTTTGAGATTATATGCAATCGAAAGTGCAGGAAACATGCC 300
QY 2774 AACTCAATCCCTCTTAATGTACATGGATGGCCAGAAAGTATGGGAGCTCTCTTGGCAG 2833
|||
DB 301 AACTCAATCCCTCTTAATGTACATGGATGGCCAGAAAGTATGGGAGCTCTCTTGGCAG 360
QY 2834 TCCGATGGAGATGGAGGATGCTTGTCAATGAAAGGACCGCTGTGTTCCATTCCGAGC 2893
|||
DB 361 TCCGATGGAGATGGAGGATGCTTGTCAATGAAAGGACCGCTGTGTTCCATTCCGAGC 420
QY 2894 TACACAAGAAAAAATGTCTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCAATGTTCTG 2953
|||
DB 421 TACACAAGAAAAAATGTCTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCAATGTTCTG 480
QY 2954 CAGCCAGACTTTCACACATTCAGAGACCTTAAATAACATGCTTAAATGCAACACCGGCC 3013
|||
DB 481 CAGCCAGACTTTCACACATTCAGAGACCTTAAATAACATGCTTAAATGCAACACCGGCC 540
QY 3014 TACCCTCTGTGAACACGAGTCTTCTCGGTGTAAGCAGAGTATCTCAGTCCGCTTGATAA 3073
|||
DB 541 TACCCTCTGTGAACACGAGTCTTCTCGGTGTAAGCAGAGTATCTCAGTCCGCTTGATAA 600
QY 3074 AAGTCAAGTCGGAACAGAACCTCCCAAGGAAAGAAATTCGAAGGAAAAATGAATTTAGCTG 3133
|||
DB 601 AAGTCAAGTCGGAACAGAACCTCCCAAGGAAAGAAATTCGAAGGAAAAATGAATTTAGCTG 660
QY 3134 TAGGTATGTGGCAGACATTTAGAGTCGCTTTGATGTTGATGATCCATGAGAACACA 3193
|||
DB 661 TAGGTATGTGGCAGACATTTAGAGTCGCTTTGATGTTGATGATCCATGAGAACACA 720
QY 3194 CAAAGATCTTTTCACTTACGCGGTGTAACATGTCGGAAGAAATTCGAAGGAGCTTTGGTT 3253
|||
DB 721 CAAAGATCTTTTCACTTACGCGGTGTAACATGTCGGAAGAAATTCGAAGGAGCTTTGGTT 780
QY 3254 TCTTAAAAATCATATCGGACACATAATGCAAAATCGGGGGCCAGAAAGCAAACTGACAGCA 3313
|||
DB 781 TCTTAAAAATCATATCGGACACATAATGCAAAATCGGGGGCCAGAAAGCAAACTGACAGCA 840
QY 3314 AGGCTTGGAGAGTAGTCCAGAACGATCAACAGAGTGTCTCAGGTGCGACCGGCCGAGAG 3373
|||
DB 841 AGGCTTGGAGAGTAGTCCAGAACGATCAACAGAGTGTCTCAGGTGCGACCGGCCGAGAG 900
QY 3374 CATCTCTCTCTTACAAATCTGATGCTTGTGCTTCTTATTTCCAAATAAAGAAAG 3433
|||
DB 901 CATCTCTCTCTTACAAATCTGATGCTTGTGCTTCTTATTTCCAAATAAAGAAAG 960
QY 3434 TCTAATTGAGCACCGCAAGGTGACACCAAAAAAACTGCTTTCGGTACCAGCAGCGCGCA 3493

Db	961	TCTAATGAGCACCAGGTCACACACCAAAAACCTGTTTCGGTACACAGCAGCGCA	1020	PR	16-MAR-2000;	2000US-0189874P.
				PR	17-MAR-2000;	2000US-0190076P.
				PR	18-APR-2000;	2000US-0198123P.
QY	3494	GACAGACTCCACAAAGAGAAATGCGTCTCGAGGGAGGACTTCCTCGCAGTTGTTCAA	3553	PR	19-MAY-2000;	2000US-0205515P.
Db	1021	GACAGACTCCACAAAGAGAAATGCGTCTCGAGGGAGGACTTCCTCGCAGTTGTTCAA	1080	PR	07-JUN-2000;	2000US-0209467P.
				PR	28-JUN-2000;	2000US-0214886P.
QY	3554	CTTGAGACCAAAATCTCACCCCTGAAACGGGGAGAGCCTGTCCAGATGCATCCCTCAGCT	3613	PR	30-JUN-2000;	2000US-0215135P.
				PR	07-JUL-2000;	2000US-0216647P.
Db	1081	CTTGAGACCAAAATCTCACCCCTGAAACGGGGAGAGCCTGTCCAGATGCATCCCTCAGCT	1140	PR	07-JUL-2000;	2000US-0216880P.
				PR	11-JUL-2000;	2000US-0217487P.
QY	3614	CGATCCGTTTCCACACCTTCAGGCTTGGCAGCTGGCTACCAAGGAAAGTTGCCATTG	3673	PR	11-JUL-2000;	2000US-0217496P.
Db	1141	CGATCCGTTTCCACACCTTCAGGCTTGGCAGCTGGCTACCAAGGAAAGTTGCCATTG	1200	PR	14-JUL-2000;	2000US-0218290P.
QY	3674	CCAAGAAGTGAAGGAATCGGGCAAGAGGGAGCACCGAACACGAGTTCGAGTTCGGA	3733	PR	26-JUL-2000;	2000US-0220963P.
Db	1201	CCAAGAAGTGAAGGAATCGGGCAAGAGGGAGCACCGAACACGAGTTCGAGTTCGGA	1260	PR	26-JUL-2000;	2000US-0220964P.
QY	3734	GAGGAGCTTGAGAAACAATAAGGGCAGTTTGTGCAGGCCCTCTGCAGAGAGAAAGNA	3793	PR	14-AUG-2000;	2000US-0224518P.
Db	1261	GAGGAGCTTGAGAAACAATAAGGGCAGTTTGTGCAGGCCCTCTGCAGAGAGAGAA	1320	PR	14-AUG-2000;	2000US-0224519P.
QY	3794	GTGCAAACTCCACGGGAAAGCGCCCTCCGTGACGCGGATCCCAAGTTACCCAGTAG	3853	PR	14-AUG-2000;	2000US-0225213P.
Db	1321	GTGCAAACTCCACGGGAAAGCGCCCTCCGTGACGCGGATCCCAAGTTACCCAGTAG	1380	PR	14-AUG-2000;	2000US-0225214P.
QY	3854	CAAGGAGAGCCCACTCACTGTCGAGTGGCGCAAAAGCTTTTCAGAACCTACCAAGCT	3913	PR	14-AUG-2000;	2000US-0225266P.
Db	1381	CAAGGAGAGCCCACTCACTGTCGAGTGGCGCAAAAGCTTTTCAGAACCTACCAAGCT	1440	PR	14-AUG-2000;	2000US-0225267P.
QY	3914	GGTCTTGCACTCCAGGCTCCACAGNAGACCGGAGGGCGGGAGTTCGCCACCAT	3973	PR	14-AUG-2000;	2000US-0225268P.
Db	1441	GGTCTTGCACTCCAGGCTCCACAGNAGACCGGAGGGCGGGAGTTCGCCACCAT	1500	PR	14-AUG-2000;	2000US-0225270P.
QY	3974	GTCTGTGGACGGGAGCGACCGGGGACGTGTTCTCTGACCTCGCGGCCCTCTGGATGA	4033	PR	14-AUG-2000;	2000US-0225477P.
Db	1501	GTCTGTGGACGGGAGCGACCGGGGACGTGTTCTCTGACCTCGCGGCCCTCTGGATGA	1560	PR	14-AUG-2000;	2000US-0225757P.
QY	4034	AAATGAGCGGTGGATTCGAGGGGAAGTGGTTCGAAACGGAATCTGAGGATGGGCTTCC	4093	PR	14-AUG-2000;	2000US-0225758P.
Db	1561	AAATGAGCGGTGGATTCGAGGGGAAGTGGTTCGAAACGGAATCTGAGGATGGGCTTCC	1620	PR	14-AUG-2000;	2000US-0225759P.
QY	4094	CGAAGGATCCATCTGGGTAA	4114	PR	18-AUG-2000;	2000US-0226279P.
Db	1621	CGAAGGATCCATCTGGGTAA	1641	PR	22-AUG-2000;	2000US-0226681P.
RESULT 9					22-AUG-2000;	2000US-0226682P.
AAK61822					23-AUG-2000;	2000US-0227009P.
ID	AAK61822 standard; cDNA; 1797 BP.				30-AUG-2000;	2000US-0228924P.
XX					01-SEP-2000;	2000US-0229287P.
AC					01-SEP-2000;	2000US-0229344P.
XX					01-SEP-2000;	2000US-0229345P.
DT	(first entry)				05-SEP-2000;	2000US-0229513P.
XX					06-SEP-2000;	2000US-0230437P.
DE	Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:6882.				06-SEP-2000;	2000US-0230438P.
XX					08-SEP-2000;	2000US-0231242P.
KW	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;				08-SEP-2000;	2000US-0231243P.
KW	cytostatic; gene therapy; vaccine; metastasis; ss.				08-SEP-2000;	2000US-0231244P.
XX					08-SEP-2000;	2000US-0231413P.
OS	Homo sapiens.				08-SEP-2000;	2000US-0231414P.
XX					08-SEP-2000;	2000US-0232080P.
FN	WO200157182-A2.				12-SEP-2000;	2000US-0231968P.
PD					14-SEP-2000;	2000US-0232398P.
XX	09-AUG-2001.				14-SEP-2000;	2000US-0232399P.
XX	17-JAN-2001; 2001WO-US001354.				14-SEP-2000;	2000US-0232401P.
XX					14-SEP-2000;	2000US-0233063P.
PR	31-JAN-2000; 2000US-0179065P.				14-SEP-2000;	2000US-0233064P.
PR	04-FEB-2000; 2000US-0180628P.				21-SEP-2000;	2000US-0234223P.
PR	24-FEB-2000; 2000US-0184664P.				21-SEP-2000;	2000US-0234274P.
PR	02-MAR-2000; 2000US-0186350P.				25-SEP-2000;	2000US-0234997P.
					25-SEP-2000;	2000US-0234998P.
					26-SEP-2000;	2000US-0235484P.
					27-SEP-2000;	2000US-0235834P.
					27-SEP-2000;	2000US-0235835P.
					29-SEP-2000;	2000US-0236327P.
					29-SEP-2000;	2000US-0236367P.
					29-SEP-2000;	2000US-0236368P.
					29-SEP-2000;	2000US-0236369P.
					29-SEP-2000;	2000US-0236370P.
					02-OCT-2000;	2000US-0236802P.
					02-OCT-2000;	2000US-0237037P.
					02-OCT-2000;	2000US-0237038P.
					02-OCT-2000;	2000US-0237039P.
					02-OCT-2000;	2000US-0237040P.

Db 747 GCTGCTTAAAAACAGTTGAAAAACCCCTGATGCGCAGCCGCTAAATTGACACAGCAATGATG 806
QY 5160 GGAAACATGAATCGGTCTTTAGGGAAGCATCTGTCAAAGTGGTCTCTTGGTTAAAAACAAGTGC 5219
Db 807 GGAAACATGAATCGGTCTTTAGGGAAGCATCTGTCAAAGTGGTCTCTCGTTTAAAAACAAGTGC 866
QY 5220 CTCCTCTCTCAGTGTCACTGATGTTGTGTGCTGTGAATCTTCGGAAGAACTGGGTGTATGA 5279
Db 867 CTCCTCTCTCAGTGTCACTGATGTTGTGTGCTGTGAATCTTCGGAAGAACTGGGTGTATGA 926
QY 5280 GACCCACAGATGAATTTGCCACACAGATGATTGGACTCTTCCTTCACCTGCTCTTCAGCC 5339
Db 927 GACCCACAGATGAATTTGCCACACAGATGATTGGACTCTTCCTTCACCTGCTCTTCAGCC 986
QY 5340 AGTCCAGTTCCTTTCTGATCATGTGATGTGACGTGAGAACTGTAGTCTGTATATCAAA 5399
Db 987 AGTCCAGTTCCTTTCTGATCATGTGATGTGACGTGAGAACTGTAGTCTGTATATCAAA 1046
QY 5400 CTTTAGAATGTTTTGAGTTTCTCGGACACAGGAACCCAGCACTTAGCATACTACAAA 5459
Db 1047 CTTTAGAATGTTTTGAGTTTCTCGGACACAGGAACCCAGCACTTAGCATACTACAAA 1106
QY 5460 TCTAATGTCTTAATGCAATCAATAAAAGAGGCTTTAAACACAGACTCCAGTTAGTAACT 5519
Db 1107 TCTAATGTCTTAATGCAATCAATAAAAGAGGCTTTAAACACAGACTCCAGTTAGTAACT 1166
QY 5520 GGTTCCTGCTAGTGGCGGTACTGTTGACAGGGCCCTGTGAGATGCCCAAGTTCCTCGAAA 5579
Db 1167 GGTTCCTGCTAGTGGCGGTACTGNTGACAGGGCCCTGTGAGATGCCCAAGTTCCTCGAAA 1226
QY 5580 GAAATGAAAAGCCAGTTACCGGTAGTGGTGTGGAACAATGGGCTAGATCATCAGCA 5639
Db 1227 GAAATGAAAAGCCAGTTACCGGTAGTGGTGTGGAACAATGGGCTAGATCATCAGCA 1286
QY 5640 GGACAGAATGCTGCTGGTGGGAGCACCCAGCTGGGCTGAGTTCTGTTCTTAC 5699
Db 1287 GGACAGAATGCTGCTGGTGGGAGCACCCAGCTGGGCTGAGTTCTGTTCTTAC 1346
QY 5700 CACTCGTGTGTTTGTGACCAATATGAGTTGCTTTAAACCTTTCTTGTCTACTATTTCCT 5759
Db 1347 CACTCGTGTGTTTGTGACCAATATGAGTTGCTTTAAACCTTTCTTGTCTACTATTTCCT 1406
QY 5760 GTTTCGAAAATGGTTCATTGACCCCTGCTTCCACCTCCACAGGACAATTTCAACAGCT 5819
Db 1407 GTTTCGAAAATGGTTCATTGACCCCTGCTTCCACCTCCACAGGACAATTTCAACAGCT 1466
QY 5820 ATTTGTAAAAAGATCACAGTCTCTTTAAAAAATAAATCTGTAAGTCAGAGGTGATGCTTG 5879
Db 1467 ATTTGTAAAAAGATCACAGTCTCTTTAAAAAATAAATCTGTAAGTCAGAGGTGATGCTTG 1526
QY 5880 AAAGAGCAGGAACCGAGTATGATGTGGAATGTCATGTCCTTTGTTCTAAAGAAAAGGCAT 5939
Db 1527 AAAGAGCAGGAACCGAGTATGATGTGGAATGTCATGTCCTTTGTTCTAAAGAAAAGGCAT 1586
QY 5940 TTTATAGCTTTTGGATATGCGCAACATACCATTAATCTGCACATAGTTGGGAGTCG 5999
Db 1587 TTTATAGCTTTTGGATATGCGCAACATACCATTAATCTGCACATAGTTGGGAGTCG 1646
QY 6000 GAAATGCAACAAACGCCAGTTTATAAACCCAGCTAGTTGGGCTATGATTGTAAGAAAAA 6059
Db 1647 GAAATGCAACAAACGCCAGTTTATAAACCCAGCTAGTTGGGCTATGATTGTAAGAAAAA 1706
QY 6060 AA 6061
Db 1707 AA 1708

RESULT 10
ACN44987
ID ACN44987 standard; cDNA; 5252 BP.
XX
AC ACN44987;
XX

DT 18-NOV-2004 (first entry)
XX Human mRNA sequence hCT1950762.
DE XX
KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003073826-A2.
XX
PD 12-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-328604/31.
XX
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
PS Claim 1; SEQ ID NO 1709; Opp; English.
XX
CC The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA gene
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published
XX
SQ Sequence 5252 BP; 1507 A; 1101 C; 1177 G; 1467 T; 0 U; 0 Other;
Query Match 15.6%; Score 1559.6; DB 11; Length 5252;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1565; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 8286 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAAGACATCTCT 8345
Db 1506 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAAGACATCTCT 1565
QY 8346 GAGGTATCATCTGGAGAGACATCACAGGAAAAACAAACCCGATGTTGCTGTAAGTCAA 8405
Db 1566 GAGGTATCATCTGGAGAGACATCACAGGAAAAACAAACCCGATGTTGCTGTAAGTCAA 1625
QY 8406 GAACGATGTTAAAAATCAGGACATGGAAGATGCACTATTAAACCGCTGACAGTGGCAAC 8465
Db 1626 GAACGATGTTAAAAATCAGGACATGGAAGATGCACTATTAAACCGCTGACAGTGGCAAC 1685
QY 8466 CAAAAATTTGAAAAGATTTTTTGATGTTGCCAAAGATGTTACAGGAGTCCACCTGCAAA 8525
Db 1686 CAAAAATTTGAAAAGATTTTTTGATGTTGCCAAAGATGTTACAGGAGTCCACCTGCAAA 1745
QY 8526 GCAGCTTAAGGAGATGCTTCTGTTTTTTCAGAAATGTTCTGGGACAGCGTGTCTCTCACC 8585
Db 1746 GCAGCTTAAGGAGATGCTTCTGTTTTTCAGAAATGTTCTGGGACAGCGTGTCTCTCACC 1805
QY 8586 AGCACACAAAGATCTCAGGATTTCCATAAAATCGAGCTGATGACAGTGTGATAAAGT 8645
Db 1806 AGCACACAAAGATCTCAGGATTTCCATAAAATCGAGCTGATGACAGTGTGATAAAGT 1865
QY 8646 GAAATAAAACCCCTACCCTGCTTACCTGGACCTGTTTAAAAAAGAGATCAGCAGTTGAAAC 8705

DR WPI; 1998-110587/10.
XX New sequences from the 20q13 amplicon - used for detecting chromosomal
PT abnormalities, particularly tumours, and for developing products for
PT treating diseases.
XX
PS Claim 1; Page 61-62; 91pp; English.
XX
CC The sequence is that of a cDNA sequence cc49, which was isolated from the
CC 20q13 amplicon. It shows homology to C2H2 zinc finger genes. It can be
CC used as a probe for the detection of chromosomal abnormalities at 20q13.
CC It and other sequences isolated from the 20q13 amplicon are consistently
CC amplified in primary tumours. These sequences are useful as probes or as
CC probe targets for monitoring the relative copy number of corresponding
CC sequences from a biological sample such as tumour cells. The sequences
CC can also be used in therapeutic applications for modulating the
CC expression of the endogenous gene or the activity of the gene product.
CC Examples of therapeutic approaches include antisense inhibition of gene
CC expression, gene therapy, and monoclonal antibodies that specifically
CC bind the gene products. The products can also be used in the treatment of
CC other diseases, e.g. age-related macular degeneration, Leber's congenital
CC amaurosis and retinitis pigmentosa
XX
SQ Sequence 1507 BP; 445 A; 354 C; 364 G; 340 T; 0 U; 4 Other;
Query Match 13.9%; Score 1387.2; DB 2; Length 1507;
Best Local Similarity 97.3%; Pred. No. 2.2e-294;
Matches 1470; Conservative 3; Mismatches 32; Indels 6; Gaps 6;
2424 CAGGTTGCTGGGATGACCTTTCTGCTCAATTTGAACACTCACTCAATTTGAAGAGC 61
2484 ACTAATGCTTTGCTGATTCATATTTGAATCAGGCAATGGAAACCTGTATGCCTTGT 2543
62 ACTAATGCTTTGCTGATTCATATTTGAATCAGGCAATGGAAACCTGTATGCCTTGT 121
2544 TTGTGGAAGAACCAAGTGACACCATCACTGAGCTTCTTAAAGTTGGAAGAGTTAGAG 2603
122 TTGTGGAAGAACCAAGTGACACCATCACTGAGCTTCTTAAAGTTGGAAGAGTTAGAG 181
2604 ACTATACATTTCTTTTGAATTTTATTAATTAATTTGCTCTGTTTGGAAACCCAGG 2663
182 ACTATACATTTCTTTTGAATTTTATTAATTAATTTGCTCTG-TTTTGGAACCCAGG 240
2664 GCTGTTAGAGGGTGTGATGACAGCTTCAAGTGGCTTATTCGAATCCAGAAATTCG 2723
241 ACTGTTTGA-GGGTGAGTGACAGGCTCTTAC-AGTGGCTTTAATCCAACTCCAGAAATTCG 298
2724 CCAACCGAATTTGAGATTATATGCAATCGAAAGTGACAGGAACATGCCAATCAATCC 2783
299 CCAACCGAATTTGAGATTATATGCAATCGAAAGTGACAGGAACATGCCAATCAATCC 358
2784 CTCTTAATGTACATGATGATGAGGAGGAGTGTGAGAGCTCTCTTGGCAGTCCGATGGAG 2843
359 CTCTTAATGTACATGAT-GGCCAGAGTGTGAGGAGCTCTCTTGGCAGTCCGATGGAG 417
2844 ATGGAGATGCTTGTCAATGAAGGACCGCTGTTGTTCCATTCGAGCTACACAGAA 2903
418 ATGGA-GATGCTTGTCAATGAAGGACCGCTGTTGTTCCATTCGAGCTACACAGAA 476
2904 AAAAATGTC-ATCCAAATCGAGGGTATATGCCCTTTGGATTGCAATGTCGAGCCAGAC 2962
477 AAAAATGTCNAATCCGAATCGAGGGGAATATGCCCTTTGGATTGCAATGTCGAGCCAGAC 536
2963 CTTTACACATTCAGAGAGCCTTAATAACATGCTTTAATGCAACACCGGCTACCTCTG 3022
537 CTTTACACATTCAGAGAGCCTTAATAACATGCTTTAATGCAACACCGGCTACCTCTG 596
3023 TGAAACAGAGTCTTTCCGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGT 3082
597 TGAACACAGAGTCTTTCCGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGT 656

QY 3083 GCGAACAGAACCTCCCAAGGAAAGAAATTCAGAGAAATGAATTTAGCTGTGAGGTATG 3142
DB 657 GCGAACAGAACCTCCCAAGGAAAGAAATTCAGAGAAATGAATTTAGCTGTGAGGTATG 716
QY 3143 TGGGCGACACATTTAGAGTTCGCTTTTGTGATGTTGAGATCCACATGAGAAACACAAAGATTC 3202
DB 717 TGGGCGACACATTTAGAGTTCGCTTTTGTGATGTTGAGATCCACATGAGAAACACAAAGATTC 776
QY 3203 TTTTCACTTACGGGTGTAAACATGTCGCGAAGAAAGATTCAGAGGCTTTGGTTCTTAAAAA 3262
DB 777 TTTTCACTTACGGGTGTAAACATGTCGCGAAGAAAGATTCAGAGGCTTTGGTTCTTAAAAA 836
QY 3263 TCACATGCGGACACATAATGCGCAAAATCGGGGGCCAGAAAGCAAACTGCGAGCAAGGCTTTGGA 3322
DB 837 TCACATGCGGACACATAATGCGCAAAATCGGGGGCCAGAAAGCAAACTGCGAGCAAGGCTTTGGA 896
QY 3323 GAGTAGTCCAGCAACGATCAACGAGGTGCTTCCAGGTGACGCGGCCGAGAGCATCTCTCTC 3382
DB 897 GAGTAGTCCAGCAACGATCAACGAGGTGCTTCCAGGTGACGCGGCCGAGAGCATCTCTCTC 956
QY 3383 TCCTTACAAATCTGATGCTTTTGTGCTTCTCTTCCAAATTAAGAAAGTCTTAATGGA 3442
DB 957 TCCTTACAAATCTGATGCTTTTGTGCTTCTCTTCCAAATTAAGAAAGTCTTAATGGA 1016
QY 3443 GCACCCGAGAGTGACACACCAAAAAAATGCTTTTCGGTACCCAGCAGCGCGCAGACAGACTC 3502
DB 1017 GCACCCGAGAGTGACACACCAAAAAAATGCTTTTCGGTACCCAGCAGCGCGCAGACAGACTC 1076
QY 3503 TCCAAAGGAGGAAATGCCCTTCTCGAGGGAGGACTTCTCTGACGTTGTTCAACTTGAAGACC 3562
DB 1077 TCCAAAGGAGGAAATGCCCTTCTCGAGGGAGGACTTCTCTGACGTTGTTCAACTTGAAGACC 1136
QY 3563 AAAATCTCACTCCCTGAAAACCGGGGAAAGGCTGTGATGATCCCTCAGCTCGATCCGTT 3622
DB 1137 AAAATCTCACTCCCTGAAAACCGGGGAAAGGCTGTGATGATCCCTCAGCTCGATCCGTT 1196
QY 3623 CACCACTTCCAGGCTTGGCAGCTGCTACCAAGGAAAGTTGCCAATTTGCCAAGAAAGT 3682
DB 1197 CACCACTTCCAGGCTTGGCAGCTGCTACCAAGGAAAGTTGCCAATTTGCCAAGAAAGT 1256
QY 3683 GAAGGAATCGGGCAAGAAAGGAGGACCCGACAAACGACGATTCGAGTTCCGAGAGGAGCT 3742
DB 1257 GAAGGAATTTGGGCAAGAAAGGAGGACCCGACAAACGACGATTCGAGTTCCGAGAGGAGCT 1316
QY 3743 TGGAGAAACAAATAAGGAGGAGTTGTCAGGCTCTCGAAGAGAAAGAGAGTGCAGAA 3802
DB 1317 TGGAGAAACAAATAAGAACCAATTTGTGAGGCTCTCGAAGAGAAAGAGAGTGCAGAA 1376
QY 3803 CTCCACGCGGAGAGCGCCCTCCGTGACGCGGATCCCAAGTTTACCCAGTAGCAAGAGAA 3862
DB 1377 CTCCACGCGGAGAGCGCCCTCCGTGACGCGGATCCCAAGTTTACCCAGTAGCAAGAGAA 1436
QY 3863 GCCCACTCACTGCTCCGAGTGGGCAAGCTTTTCAAGAACTTACCAGCTGGTCTTTGCA 3922
DB 1437 GCCCACTCACTGCTCCGAGTGGGCAAGCTTTTCAAGAACTTACCAGCTGGTCTTTGCA 1496
QY 3923 CTCACGGGTCC 3933
DB 1497 CTCACGGGTCC 1507
RESULT 13
ADP03055/c
ID ADP03055 standard; cDNA; 124987 BP.
XX
AC ADP03055;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human housekeeping gene cdNA #96.
XX
KW ss; gene; human; housekeeping gene; capture probe; DNA microarray;
cancer.

XX Homo sapiens.
OS JP2004135552-A.
PN 13-MAY-2004.
XX 16-OCT-2002; 2002JP-00302278.
XX 16-OCT-2002; 2002JP-00302278.
XX (NIGA) NGK INSULATORS LTD.
XX WPI; 2004-404613/38.
XX Novel human housekeeping gene commonly expressed in different human
PT tissues, useful as capture probe in DNA microarray.
XX Claim 1; SEQ ID NO 96; 940pp; Japanese.
XX The invention relates to a human housekeeping gene. The housekeeping gene
CC is useful as a capture probe in DNA microarray. The housekeeping gene is
CC useful as a reference with respect to the normal gene expression measured
CC on various conditions, and an expression database of normal human tissue
CC is useful as reference information for using a disease related gene for
CC diagnosis or treatment. The housekeeping gene when used as capture probe,
CC enables diagnosis of cancer. The present sequence represents a human
CC housekeeping gene cDNA.
XX Sequence 124987 BP; 31878 A; 32311 C; 31059 G; 29739 T; 0 U; 0 Other;
SQ
Query Match 10.1%; Score 1006.8; DB 12; Length 124987;
Best Local Similarity 95.5%; Pred. No. 5.9e-210;
Matches 1083; Conservative 0; Mismatches 37; Indels 14; Gaps 4;
QY 1 CCATCATATTTCTTATTTTGGGGGAGAGGGG-----AGACTTGCTCTGTGGCCCA 54
DB 1128 CCATCATATTTCTTATTTTGGGCAGAGGGGGGANGTCAAGTCTCGCTCTGTGGCCCA 1069
QY 55 GCCTGGA--CCAGTGGTGGATCTGGGTCATGCAACCTCCACCTCTCGGTTCAAGTG 112
DB 1068 GGCTGGAATGCAGTGGCGGATCTTGGCTCACTGCAACCTCCACCTCTCGGTTCAAGTG 1009
QY 113 ATCCCAATAGCTGGATTACAGGTGTATTACCATGCCCGAGCTTAATTTTGTATTTT 172
DB 1008 ATTCCCAAGTAGCTGGATTACAGGTGTATTACCATGCCCGAGCTTAATTTTGTATTTT 949
QY 173 TAGCAGATAAGGGTTTACCATGTTGGCAGGCTGGTCTCCAACTCTCTGGCTCATGTG 232
DB 948 TAGCAGATAAGGGTTTACCATGTTGGTCAAGCTGGTCTCCAACTCTGACCTCAAGTG 889
QY 233 ATCCACCCACTTCGGCTTCCCAAGCATTTGGAGTATAGGTGTGAGCCACTATACCCGTC 292
DB 888 ATCCACCCACTTCGGCTTCCCAAGCATTTGGAGTATAGGTGTGAGCCACTATACCCGTC 829
QY 293 CTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTCTAAGGATGT 352
DB 828 CTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTGGAAAGGATGT 769
QY 353 CAGTAGAGAAGTGGAGTCTCCCAAAATTACAGTTTCAAGTATTAGTCAAGTTTCTAAAT 412
DB 768 CAGTAGAGAAGTGGAGTCTCCCAAAATTACAGTTTCAAGTATTAGTCAAGTTTCTAAAT 709
QY 413 ACAGTAATAATGTGTAGAGCTGACATAGGACTTAAGTGTGTTTGTGTTTGTGTTTGTGTTT 472
DB 708 ACAGTAATAATGTGTAGAGCTGACATAGGACTTAAGTGTGTTTGTGTTTGTGTTTGTGTTT 654
QY 473 TTTCAAAATCTCACTCAACTTGTATTTTGTAAATAGGACATTAAGGACATTAAGGACATTA 532
DB 653 TTTCAAAATCTCACTCAACTTGTATTTTGTAAATAGGACATTAAGGACATTAAGGACATTA 595
QY 533 AACTCCACTATTGCGCTATTGGCCACTATTGATTTTAAAAAATAAGCGTATTATTAGCAT 592

DB 594 AACTCCACTATTGCGCATATTGCCATATTATTGATTTTTTAAAAAATAAGCGTATTATTAGCAT 535
QY 593 CTAAAGTAGAAGGACCTCAATAAATGAGTCTTTGTTCTTTGGCCAGGGAACAGCGT 652
DB 534 CTGCAAGTAGAAGGCGCTCAATAAATGAGTCTTTGTTCTTTGGCCAGGGAACAGCGT 475
QY 653 TGTGAGAATTTGATAAAGTCTTTTCTAGGGTATGTGCTGTTATTTCAGTATAAAGCTTTGCC 712
DB 474 TGTGAGCAATTTGTTAACTGTTTCTAGGGTATGTGCTGTTATTTCAGTATAAAGCTTTGCC 415
QY 713 TGGGACGCTAGCAATTCAGTAAATCTTGTGAATAAGCAAAATGAACTTAAGCTTTCTATG 772
DB 414 TGGGACGCTAGCAATTCAGTAAATCTTGTGAATAAGCAAAATGAACTTAAGCTTTCTATG 355
QY 773 TATAGAAACCTAAGTCACTTCACTTCTGATTAGCAGAGTAAATTCGAATATTTCTTTCAAT 832
DB 354 TATAGAAACCTAAGTCACTTCACTTCTGATTAGCAGAGTAAATTCGAATATTTCTTTTCAGT 295
QY 833 GTGTAGCTCTATCCCGAGAACCCACAGATATTGGAACTGTAAAGGCCATCTTATAGTTTA 892
DB 294 GTGTAGATCTATCCCGAGAACCCACAGATATTGGAACTGTAAAGGCCATCTTATAGTTTA 235
QY 893 ACCAACTCGTTAAATAGATATAAGAAAGATGTGGTATGTGGCAGTGAACAATCGAAGT 952
DB 234 ACCAACTCGTTAAATAGATATAAGAAAGATGTGGTATGTGGCAGTGAACAATCGAAGT 175
QY 953 TGTGACTAGAACTCGGGTCTCTGGAGTGTCTTATATATATCACCAAGCTGGTCAACG 1012
DB 174 TGTGACTAGAACTCGGGTCTCTGGAGTGTCTTATATATATCACCAAGCTGGTCAACG 115
QY 1013 CCATGTTGTTGATCTCTCCATTTGTGATAGCAACAAGAAAGACTTCAGGACATTTCTTTCT 1072
DB 114 CAGTGTGTTGATCTCTCCATTTGTGATAGCAACAAGAAAGACTTCAGGACATTTCTTTCT 55
QY 1073 TTACCTTAATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATC 1126
DB 54 TTACCTTAATCCTTGATCTGCAGTCTTATTTAGAAAAGCTTAATGTTAAAGATC 1
RESULT 14
ADS88553/c
ID ADS88553 standard; cDNA; 124990 BP.
XX ADS88553;
XX 18-NOV-2004 (first entry)
XX Human housekeeping gene cDNA sequence SEQ ID NO:96.
XX housekeeping gene; human; probe; hybridisation; DNA microarray;
KW disease investigation; disease diagnosis; gene; ss.
XX Homo sapiens.
XX WO2004035785-A1.
XX 29-APR-2004.
XX 16-OCT-2002; 2002WO-JP010753.
XX 16-OCT-2002; 2002WO-JP010753.
XX (NIGA) NGK INSULATORS LTD.
XX Aburatani H, Yamamoto S;
XX WPI; 2004-357219/33.
XX Human housekeeping and tissue-specific genes, gene sets and transcription
PT products, probes and microarrays derived from them for reference use in
PT investigation and diagnosis of disease.
XX Claim 5; SEQ ID NO 96; 372pp; Japanese.

XX The present invention describes 1189 housekeeping genes and gene sets
CC containing them, expressed in 35 different human tissues. Also described:
CC (1) RNA and cDNA transcription products of the housekeeping genes; (2)
CC oligonucleotide probes hybridizing to the housekeeping genes; (3) DNA
CC microarrays containing the oligonucleotide probes; (4) 1704 genes
CC expressed specifically in particular tissues; (5) RNA and cDNA
CC transcription products of the tissue-specific genes; (6) oligonucleotide
CC probes hybridizing to the tissue-specific genes; and (7) DNA microarrays
CC containing the oligonucleotide probes. The tissue-specific genes are
CC expressed specifically in one of brain, thymus, pituitary, thyroid,
CC trachea, lung, breast, skin, skeletal muscle, heart, liver, spleen,
CC kidney, adrenal gland, pancreas, stomach, small intestine, colon,
CC bladder, prostate gland, testis, ovary, placenta, uterus, bone marrow,
CC foetal brain or foetal liver. The housekeeping genes can be used as
CC reference genes for the investigation and diagnosis of disease. The
CC present sequence represents a human housekeeping gene cDNA sequence which
CC is given in the exemplification of the present invention.

XX
SQ Sequence 124990 BP; 31879 A; 32311 C; 31060 G; 29740 T; 0 U; 0 Other;
Query Match 10.1%; Score 1006.8; DB 13; Length 124990;
Best Local Similarity 95.5%; Pred. No. 5.9e-210;
Matches 1083; Conservative 0; Mismatches 37; Indels 14; Gaps 4;

QY 1 CCATCATATTTCTTTATTTTGGGGGAGAGGGG-----AGACTTGTCTCTGTGCCCA 54
DB 1128 CCATCATATTTCTTTTGGGCAAGAGGGGGAGTCAAGTCTCGCTCTGTGCCCA 1069

QY 55 GCGTGA--CCAGTGGTGCATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTG 112
DB 1068 GCGTGAATGCACTGGCGCATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTG 1009

QY 113 ATTCCCAATAGCTGGATACAGGTGTGTTACCATGCCAGCTAAATTTTGTATTTT 172
DB 1008 ATTCCCAAGTAGCTGGATACAGGTGTGTTACCATGCCAGCTAAATTTTGTATTTT 949

QY 173 TAGCAGATAAGGGTTTCCACCATGTTGGCCAGGCTGGTCTCCAACTCCTGGCCTCATGTG 232
DB 948 TAGCAGATAAGGGTTTCCACCATGTTGGTCAGGCTGGTCTCCAACTCCTGGCCTCATGTG 889

QY 233 ATCCACCCACTTGGCTTCCCAAGCAATGGGAGTATAGGTGTGAGCCACTATACCCGTC 292
DB 888 ATCCACCCACTTGGCTTCCCAAGCAATGGGAGTATAGGTGTGAGCCACTATGCGCGC 829

QY 293 CTCACATCATATTTCTAAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTCTAAAGCATGT 352
DB 828 CTCACATCATATTTCTAAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTGGAAAGGATGT 769

QY 353 CAGTAGAGAAGTGGAGTCTCCCAAAATACAGTTTCCACGTATTAGTCAAGTTTCTAAAT 412
DB 768 CAGTAGAGAAGTGGAGTCTCCCAAAATACAGTTTCCACATATTAGTCAAGTTTCTAAAT 709

QY 413 ACAGTAATATGTGTAGAGCTGACATAGGACTTAACCTGGTTTTTTTTTTTTTTTTTTT 472
DB 708 ACAGTAATATGTGTAGAGCTGACATAGGACTTAACCTTAG-----TTTTTGTTTTTTTT 654

QY 473 TTTCAAATCTCACTGAACCTTTGATTTTCTTAATAAGGACATTAAAAAACCACAAA 532
DB 653 TTTCAAATCTCACTGAACCTTTGATTTTCTTAATAAGGACATT-AAAAAACCACAAA 595

QY 533 AACTCCACTATTGCGCTATTGCGCACTATTGATTTTTTAAAAAATAGCGTATTTTAGCAT 592
DB 594 AACTCCACTATTGCGCTATTGCGCACTATTGATTTTTTAAAAAATAGCGTATTTTAGCAT 535

QY 593 CTAAAGTAGGAAGGACCTCAATAAATAGCTTTTGTCTTGGCCAGGAAAAACAGCGT 652
DB 534 CTGCAAGTAGGAAGGCGCTCAATAAATAGCTTTTGTCTTGGCCAGGAAAAACAGCGT 475

QY 653 TGTCAAGTTTGAATTAAGTCTTTTCTAGGATATGCTGTTATTAGTTTAAACCTTGGC 712
DB 474 TGTCAAGTTTGAATTAAGTCTTTTCTAGGATATGCTGTTATTAGTTTAAACCTTGGC 415

QY 713 TGGAGCGCTAGCATTCAGTAAATACTGTTGAATAAGCAATGAACTTAAGCTTCTATG 772
DB 414 TGGAGCGCTAGCATTCAGTAAATACTGTTGAATAAGCAATGAACTTAAGCTTCTATG 355

QY 773 TATAGAAACCTAAGTCACTTCACTTCTGATTAGCAGAGTAATTGAAATATTTTCAAT 832
DB 354 TATAGAAACCTAAGTCACTTCACTTCTGATTAGCAGAGTAATTGAAATATTTTCAAT 295

QY 833 GTCTAGCTCTATCCCGAGACCAAGATATTCGAACTGTAAGGCCATCTCTATAGTTA 892
DB 294 GTCTAGCTCTATCCCGAGACCAAGATATTCGAACTGTAAGGCCATCTCTAGAGTTA 235

QY 893 ACCAACTGCGTTAAATAGATAAATAGAAAGATGTGTTATGTGGCAGTCACTTGAAGGT 952
DB 234 ACCAACTGCGTTAAATAGATAAATAGAAAGATGTGTTATGTGGCAGTCACTTGAAGGT 175

QY 953 TGTGACTAGAACTCGGGTCTCTCGAGTGTCTTATTATATCACCAAGCTGGTCAACG 1012
DB 174 TGTGACTAGAACTCGGGTCTCTCGAGTGTCTTATTATATCACCAAGCTGGTCAACG 115

QY 1013 CCATGTTGATCCTCCATTTGTGATAGCAACAAAGAAAGACTTCAGGACATTTCTTCT 1072
DB 114 CCATGTTGATCCTCCATTTGTGATAGCAACAAAGAAAGACTTCAGGACATTTCTTCT 55

QY 1073 TTACCCTAATCTCTGATCTGAGTCTTATTAGAAAAGCTTAATGTTAAAGATC 1126
DB 54 TTACCCTAATCTCTGATCTGAGTCTTATTAGAAAAGCTTAATGTTAAAGATC 1

RESULT 15
ACN44984

ID ACN44984 standard; DNA; 26345 BP.

XX ACN44984;

XX 18-NOV-2004 (first entry)

XX Mouse genomic sequence MCG6549.

XX Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.

XX Mus musculus.

XX WO2003073826-A2.

XX 12-SEP-2003.

XX 28-FEB-2003; 2003WO-US006235.

XX 01-MAR-2002; 2002US-00087192.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW;

XX WPI; 2003-328604/31.

XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma comprises a nucleotide sequence.

XX Claim 1; SEQ ID NO 1705; Opp; English.

XX The present invention relates to novel DNA and protein sequences which are associated with carcinomas. The sequences are useful for: (i) for screening drug candidates; (ii) for screening of bioactive agent capable of binding to Carcinoma Associated Protein (CAP); (iii) for screening of a bioactive agent capable of modulating the activity of CAP; (iv) for evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining Carcinoma Associated (CA) gene copy number. In addition, the CA genes are useful as DNA vaccines and the CAP are useful as markers of

Qy 4114 AG 4115
||
Db 11391 AG 11392

Search completed: July 31, 2005, 12:46:01
Job time : 3149 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 10:56:42 ; Search time 27398 Seconds
(without alignments)
17685.676 Million cell updates/sec

Title: us-08-731-499-9_COPY_1_10000

Perfect score: 10000

Sequence: 1 CATCATATTTCTTATTTT.....ATTCCTTAGATACGCAGTGG 10000

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_uni.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	10000	100.0	10365	6	BD085733	Genes from the 20q13 amplicon and their uses.
2	10000	100.0	121143	9	AF312915	BD085733
3	9980.6	99.8	128871	9	AL157838	Genes from the 20q13 amplicon and their uses.
4	9943.6	99.4	105023	2	AC116668	BD085733
5	2132	21.3	165471	2	AC015713	Genes from the 20q13 amplicon and their uses.
6	2015	20.2	111123	9	AP000365	BD085733
7	2015	20.2	187590	2	AC141274	Genes from the 20q13 amplicon and their uses.
8	1639.4	16.4	5632	6	AR157086	BD085733
9	1639.4	16.4	5632	6	AF041259	Genes from the 20q13 amplicon and their uses.
10	1559.6	15.6	3186	6	C0726295	BD085733
11	1387.2	13.9	1507	6	AR070326	Genes from the 20q13 amplicon and their uses.
12	1387.2	13.9	1507	6	BD085727	Genes from the 20q13 amplicon and their uses.
13	1223.6	12.2	59586	9	HS449017	Genes from the 20q13 amplicon and their uses.
14	1223.6	12.2	59586	9	HS449017	Genes from the 20q13 amplicon and their uses.
15	1006.8	10.1	124990	2	AC099079	Genes from the 20q13 amplicon and their uses.
16	850.6	8.5	246332	2	AC099079	Genes from the 20q13 amplicon and their uses.
17	828.8	8.3	184420	2	AC118373	Genes from the 20q13 amplicon and their uses.
18	820	8.2	114285	10	AL844576	Genes from the 20q13 amplicon and their uses.
19	820	8.2	208979	2	AC023610	Genes from the 20q13 amplicon and their uses.

20	345	3.5	46985	2	AC020762	AC020762 Homo sapi
21	335.8	3.4	469	6	BD060160	BD060160 Secreted
22	274.8	2.7	1196	10	BC046393	BC046393 Mus muscu
23	244.6	2.4	251	11	G65441	G65441 stdJ434Pl.2
24	239.8	2.4	191540	9	AC144780	AC144780 Pan trogl
25	235.8	2.4	103287	2	AC084034	AC084034 Homo sapi
26	234	2.3	110000	2	BX572642	BX572642 Homo sapi
27	234	2.3	113051	9	AL353150	AL353150 Human DNA
28	233.8	2.3	172571	9	AC006064	AC006064 Homo sapi
29	233.4	2.3	83639	9	AL353787	AL353787 Human DNA
30	232.6	2.3	68620	2	AC031999	AC031999 Homo sapi
31	232.6	2.3	118695	9	HSJ672M15	AL045643 Human DNA
32	232	2.3	3512	5	BC072191	BC072191 Xenopus l
33	232	2.3	31484	9	AC131958	AC131958 Homo sapi
34	232	2.3	56330	2	AL353694	Continuation (4 of
35	231.2	2.3	177964	9	AC016705	AC016705 Homo sapi
36	231.2	2.3	198707	9	AC108451	AC108451 Homo sapi
37	231	2.3	156839	9	AC107948	AC107948 Homo sapi
38	231	2.3	166973	9	AC084117	AC084117 Homo sapi
39	230.8	2.3	110000	2	AL831785	Continuation (3 of
40	230.8	2.3	131682	9	AL672277	AL672277 Human DNA
41	230.4	2.3	2081	9	HSMB06272	BX537592 Homo sapi
42	230.2	2.3	172579	9	AC008064	AC008064 Homo sapi
43	229.8	2.3	48593	2	AC113138	AC113138 Homo sapi
44	229.8	2.3	218892	9	AC007225	AC007225 Homo sapi
45	229.6	2.3	71819	9	AC007536	AC007536 Homo sapi

ALIGNMENTS

RESULT 1	BD085733	10365 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD085733	Genes from the 20q13 amplicon and their uses.			
DEFINITION	BD085733	Genes from the 20q13 amplicon and their uses			
ACCESSION	BD085733	Genes from the 20q13 amplicon and their uses			
VERSION	BD085733.1	GI:22631343			
KEYWORDS	JP 2001524802-A/9.				
SOURCE	synthetic construct				
ORGANISM	other sequences; artificial sequences.				
REFERENCE	1 (bases 1 to 10365)				
AUTHORS	Gray,J.W., Collins,C.C., Hwang,S.I., Godfrey,T., Kowbel,D. and Rommens,J.				
TITLE	Genes from the 20q13 amplicon and their uses				
JOURNAL	Patent: JP 2001524802-A, 9 04-DEC-2001;				
COMMENT	THE REGENTS OF THE UNIVERSITY OF CALIFORNIA				
	OS Artificial Sequence				
	PN JP 2001524802-A/9				
	PD 04-DEC-2001				
	PF 15-JUL-1997 JP 1998506264				
	PR 15-JUL-1996 US 08/680395,16-OCT-1996 US 08/731499 PR				
	17-TAN-1997 US 08/785532				
	PI JOE W GRAY, COLIN CONRAD COLLINS, SOO IN HWANG, TONY GODFREY, PI DAVID KOWBEL,				
	PI JOHANNA ROMMENS				
	PC C12N15/11,C12Q1/68,A61K48/00				
	CC Description of Artificial Sequence:Genomic Sequence encoding				
	CC ZABC1				
	FH Key Location/Qualifiers				
	FT source 1..10365				
	FT /organism='Artificial Sequence'.				
FEATURES	Location/Qualifiers				
source	1..10365				
	/organism="synthetic construct"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32630"				
ORIGIN					
Query Match	100.0%;	Score	10000;	DB 6;	Length 10365;
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 10000;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	

Db	2161	AGCACTGATTTAAACTGATTTTAACTGGATGAAATTTCTGATTTAAATAAGTGTACTGACT	2222
Qy	2221	GGATAAAATGCGCAATGATTTTAATTAACAAGCACGCTTTAAACAGGATGCGCTATATATATAGT	2280
Db	2221	GGATAAAATGCGCAATGATTTTAATTAACAAGCACGCTTTAAACAGGATGCGCTATATATATAGT	2280
Qy	2281	TAAAGTGAAGCAATTTGAATTTAGGTACCTTCTCTGCTGCGTGGAAAAAGACCGTATGACTC	2340
Db	2281	TAAAGTGAAGCAATTTGAATTTAGGTACCTTCTCTGCTGCGTGGAAAAAGACCGTATGACTC	2340
Qy	2341	ACCACACACCGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTGTCTCTGTTTTTTTTTCCCTCT	2400
Db	2341	ACCACACACCGCCTTCTCTTCGCTCTGAGTGTAGCTAACCGTGTCTCTGTTTTTTTTTCCCTCT	2400
Qy	2401	AGGGTTGGAAATCCCTTGTCTCCAGGTGTCTGGGATTTGACTTCTTGCTCAATTTGAAACA	2460
Db	2401	AGGGTTGGAAATCCCTTGTCTCCAGGTGTCTGGGATTTGACTTCTTGCTCAATTTGAAACA	2460
Qy	2461	CTCAATTCAAATGGAGACAAAGAGAACTAATGCTTTTGTGCTGATTCATATTTGAAATCGAGGC	2520
Db	2461	CTCAATTCAAATGGAGACAAAGAGAACTAATGCTTTTGTGCTGATTCATATTTGAAATCGAGGC	2520
Qy	2521	ATTGGGAACCCCTGTATGCTTGTGTTGTGGAAAGAACAGTGACACCATCACTGAGCTTCC	2580
Db	2521	ATTGGGAACCCCTGTATGCTTGTGTTGTGGAAAGAACAGTGACACCATCACTGAGCTTCC	2580
Qy	2581	TAAAGTTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACCTTTATAATAATATTT	2640
Db	2581	TAAAGTTTCGAAGAAGTTAGAGGACTATACACTTTCTTTTGAACCTTTATAATAATATTT	2640
Qy	2641	TGCTCTGGTTTTTGGAAACCCAGGCGCTGTATAGAGGGTGAGTGACAAGTCTTTACAAGTGGC	2700
Db	2641	TGCTCTGGTTTTTGGAAACCCAGGCGCTGTATAGAGGGTGAGTGACAAGTCTTTACAAGTGGC	2700
Qy	2701	CTTATTTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGAGTTATATGCAATCGAAGTGA	2760
Db	2701	CTTATTTCCAACTCCAGAAATTGCCCAACGGAACTTTGAGATTTATGCAATCGAAGTGA	2760
Qy	2761	CAGAAACATGCCAACTCAATCCCCTCTTAATGTACATGGATGGGCCAAGAAGTCATTTGGCA	2820
Db	2761	CAGAAACATGCCAACTCAATCCCCTCTTAATGTACATGGATGGGCCAAGAAGTCATTTGGCA	2820
Qy	2821	GCTCTCTTGGCAGTCCGATGGAGATGAGAGATGCGCTTGTCAATGAAAGGACCGCTGTTG	2880
Db	2821	GCTCTCTTGGCAGTCCGATGGAGATGAGAGATGCGCTTGTCAATGAAAGGACCGCTGTTG	2880
Qy	2881	TTCCATTTCCGAGCTACACAGAAAAAATGTCAATCCAAATTCGAGGGGTATATGCCCTTGG	2940
Db	2881	TTCCATTTCCGAGCTACACAGAAAAAATGTCAATCCAAATTCGAGGGGTATATGCCCTTGG	2940
Qy	2941	ATTGCATGTTCTGCAGCCAGACCTTACACATTCAGAAGACCTTTAATAACATGTCTTTAA	3000
Db	2941	ATTGCATGTTCTGCAGCCAGACCTTACACATTCAGAAGACCTTTAATAACATGTCTTTAA	3000
Qy	3001	TGCAACACCGGCGCTACCCCTCTGTGAAACGACGAGTCTTTTCGGGTTTGAAGCAGAGTATCTCA	3060
Db	3001	TGCAACACCGGCGCTACCCCTCTGTGAAACGACGAGTCTTTTCGGGTTTGAAGCAGAGTATCTCA	3060
Qy	3061	GTCCGCTTGATATAAGTCAAGTGGGAACAGAACTTCCCAAGGAAAAAGATTCGAAGGAAA	3120
Db	3061	GTCCGCTTGATATAAGTCAAGTGGGAACAGAACTTCCCAAGGAAAAAGATTCGAAGGAAA	3120
Qy	3121	ATGAAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGTAGTTGAGATCC	3180
Db	3121	ATGAAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGTAGTTGAGATCC	3180
Qy	3181	ACATGAGAACACAAAAGATTTCTTTCACTTACCGGGTGTAACATGTGCGGAAGAAGATTCA	3240
Db	3181	ACATGAGAACACAAAAGATTTCTTTCACTTACCGGGTGTAACATGTGCGGAAGAAGATTCA	3240
Qy	3241	AGGAGCCTTGGTTTTCTTAAAAATTCACATGCGGACACATAATTGGCAATCGGGGGCCAGAA	3300
Db	3241	AGGAGCCTTGGTTTTCTTAAAAATTCACATGCGGACACATAATTGGCAATCGGGGGCCAGAA	3300

Qy	3301	GCAAATCTGCAGCAAGGCTTGGAGTAGTCCAGCAACGATCAACAGAGTGTCTCAGTGTC	3360
Db	3301	GCAAATCTGCAGCAAGGCTTGGAGTAGTCCAGCAACGATCAACAGAGTGTCTCAGTGTC	3360
Qy	3361	ACGGCGCGGAGAGCATCTCTCTCTTACAAAATCTGCATGGTGTGTGGCTTCTCTATTTC	3420
Db	3361	ACGGCGCGGAGAGCATCTCTCTCTTACAAAATCTGCATGGTGTGTGGCTTCTCTATTTC	3420
Qy	3421	CAAAATAAGAAAGTCTTAATTGAGCACCGCAAGGTGCAACACAAAATACTGCTTTTCGGTA	3480
Db	3421	CAAAATAAGAAAGTCTTAATTGAGCACCGCAAGGTGCAACACAAAATACTGCTTTTCGGTA	3480
Qy	3481	CCAGCAGCGCGCAGACAGACTCTTCCACAAGSAGGAATCCGCTCTCCAGAGSAGACTTCC	3540
Db	3481	CCAGCAGCGCGCAGACAGACTCTTCCACAAGAGGAATGCCGTCTCCAGAGSAGACTTCC	3540
Qy	3541	TGCAGTTGTTCAACTTGAGACCAAAATCTCACCTTGAAACGGGGAAGAAGCTGTTCAGAT	3600
Db	3541	TGCAGTTGTTCAACTTGAGACCAAAATCTCACCTTGAAACGGGGAAGAAGCTGTTCAGAT	3600
Qy	3601	GCATCCCTCAGCTCGATCCGTTTACCACACTTCCAGGCTTGGCAGCTGGCTTACCAAGGAA	3660
Db	3601	GCATCCCTCAGCTCGATCCGTTTACCACACTTCCAGGCTTGGCAGCTGGCTTACCAAGGAA	3660
Qy	3661	AAGTTGCATTTTGCACAAAGAGTGAAGGAATCGGGGCAAGAAGGAGCACCGACAACGACG	3720
Db	3661	AAGTTGCATTTTGCACAAAGAGTGAAGGAATCGGGGCAAGAAGGAGCACCGACAACGACG	3720
Qy	3721	ATTTCGAGTTCGGAAGAGGAGCTTGGAGAAACAATAAGGCGAGTTGTGCAGGCTCTTCG	3780
Db	3721	ATTTCGAGTTCGGAAGAGGAGCTTGGAGAAACAATAAGGCGAGTTGTGCAGGCTCTTCG	3780
Qy	3781	AAGAGAAGAGAAGTGCAAACACTCCACACGCGAAGCGCCCTCCGTGGACGCGGATCCCA	3840
Db	3781	AAGAGAAGAGAAGTGCAAACACTCCACACGCGAAGCGCCCTCCGTGGACGCGGATCCCA	3840
Qy	3841	AGTTACCAGTAGCAAGAGGAGGCCACTCACTGCTCCGAGTCCGGCAAAAGCTTTTCAGAA	3900
Db	3841	AGTTACCAGTAGCAAGAGGAGGCCACTCACTGCTCCGAGTCCGGCAAAAGCTTTTCAGAA	3900
Qy	3901	CCTTACCACAGCTGGTCTTTCACATCCAGGTCACCAAGAGAGCACGGAGGCGCGCGCG	3960
Db	3901	CCTTACCACAGCTGGTCTTTCACATCCAGGTCACCAAGAGAGCACGGAGGCGCGCGCG	3960
Qy	3961	AGTCGCGCCACCATGTCTGTGGACGGAGGCAGCGCGGGAGCTGTCTCTGACCTCGCCG	4020
Db	3961	AGTCGCGCCACCATGTCTGTGGACGGAGGCAGCGCGGGAGCTGTCTCTGACCTCGCCG	4020
Qy	4021	CCCTCTGAGATGAAATTGGAGCCGTGGATCGAGGGGAAGGTGTTCTGAAAGACGGATCTG	4080
Db	4021	CCCTCTGAGATGAAATTGGAGCCGTGGATCGAGGGGAAGGTGTTCTGAAAGACGGATCTG	4080
Qy	4081	AGGATGGGCTTCCGGAAGGAATCCATCTGGTAAGCTGCCCTCTCTCCGTCGGCTGCTGT	4140
Db	4081	AGGATGGGCTTCCGGAAGGAATCCATCTGGTAAGCTGCCCTCTCTCCGTCGGCTGCTGT	4140
Qy	4141	TCGCGCTGTGTGTGTCTGTCTCCCGTCTCCCGCTCTCTATTTCCCATCTCCAGACAACGC	4200
Db	4141	TCGCGCTGTGTGTGTCTGTCTCCCGTCTCCCGCTCTCTATTTCCCATCTCCAGACAACGC	4200
Qy	4201	TGGCCAGGAATGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTGGTATCACTCT	4260
Db	4201	TGGCCAGGAATGGGTTTGGAGAGCCAGAGTCAAGTCCAGGCTCTTTTGGTATCACTCT	4260
Qy	4261	GTGTAAAGTCAATTTAACTCTCAGGGCTTAAATTTTCTCATTTCTGTAATAACAGGGTTGA	4320
Db	4261	GTGTAAAGTCAATTTAACTCTCAGGGCTTAAATTTTCTCATTTCTGTAATAACAGGGTTGA	4320
Qy	4321	GTTAAGAGGTCCTCTGTTCTGAAAATAATAATAATTTTTTTAAACGTGTATCGTTTTGC	4380
Db	4321	GTTAAGAGGTCCTCTGTTCTGAAAATAATAATAATTTTTTTAAACGTGTATCGTTTTGC	4380

QY 4381 TCACAAAACACACTTTAAAAAAAATAAATACATTTGTGCATCCAGCCCAATGCACTGCTTCT 4440
DB 4381 TCACAAAACACACTTTAAAAAAAATAAATACATTTGTGCATCCAGCCCAATGCACTGCTTCT 4440
QY 4441 TAACTGGGGCGATTTTGTTCCTCAATCAGTATCTGGCAATGTCCTGGAGGCAATTTTGGTTGT 4500
DB 4441 TAACTGGGGCGATTTTGTTCCTCAATCAGTATCTGGCAATGTCCTGGAGGCAATTTTGGTTGT 4500
QY 4501 CATATCTGTGTGTGGGTGTGCTGTGTCATCCAGTGGGCGAGAGCCAGGCACTGCT 4560
DB 4501 CATATCTGTGTGTGGGTGTGCTGTGTCATCCAGTGGGCGAGAGCCAGGCACTGCT 4560
QY 4561 CAGCATGTPACAGTGCACAGGACGCCCATCATCAAGAAATTAATCTGGTCCCAATGTC 4620
DB 4561 CAGCATGTPACAGTGCACAGGACGCCCATCATCAAGAAATTAATCTGGTCCCAATGTC 4620
QY 4621 AATAGTTTGGAGCAATTTGAGAGACCTTAGCCTTCACTTAAAGTTTCTGGCGTTTCTGTGATCT 4680
DB 4621 AATAGTTTGGAGCAATTTGAGAGACCTTAGCCTTCACTTAAAGTTTCTGGCGTTTCTGTGATCT 4680
QY 4681 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAGGTACTGGGAGTGATCAACTAGAGCCAG 4740
DB 4681 TTTTCTGTAGTGAATTTCTAGTGGCCATAAAGGTACTGGGAGTGATCAACTAGAGCCAG 4740
QY 4741 GAATATTTTGGGCGAGCGTTTGGTGTCTCCAAAACCTTGTCTTCTGTCTGGCAAG 4800
DB 4741 GAATATTTTGGGCGAGCGTTTGGTGTCTCCAAAACCTTGTCTTCTGTCTGGCAAG 4800
QY 4801 CTAGTATCCATTTATAGGTACCTCAGGAACCCCAATGATTTGTCTATAAAATACAAGGAAT 4860
DB 4801 CTAGTATCCATTTATAGGTACCTCAGGAACCCCAATGATTTGTCTATAAAATACAAGGAAT 4860
QY 4861 GTGAGCACACTGAAGACATTTTAAAGAGGCTCAATTTGCTCAGCAGAAATTTTCACTGTAC 4920
DB 4861 GTGAGCACACTGAAGACATTTTAAAGAGGCTCAATTTGCTCAGCAGAAATTTTCACTGTAC 4920
QY 4921 TAGTGGCATTTATAGAAAGAGAGGTGATCACTGAAGGCATCTCACATAATTTTCTCTGA 4980
DB 4921 TAGTGGCATTTATAGAAAGAGAGGTGATCACTGAAGGCATCTCACATAATTTTCTCTGA 4980
QY 4981 GCCCTGGTGGCGTTTATCTAGGGCAAGGATTTCCACCTGTGTGTGGAGTTGGCCCATCC 5040
DB 4981 GCCCTGGTGGCGTTTATCTAGGGCAAGGATTTCCACCTGTGTGTGGAGTTGGCCCATCC 5040
QY 5041 TCATCTGTAGCCAGAGCTTCTCCTATCAGAGTTTATGTAATTTTGTGTAATAGAGATCTTG 5100
DB 5041 TCATCTGTAGCCAGAGCTTCTCCTATCAGAGTTTATGTAATTTTGTGTAATAGAGATCTTG 5100
QY 5101 CTGCTTAAACAGTTGMAAAGACCCCTGATGGCGAGCGCTAATTCACAAAGCAATGATGG 5160
DB 5101 CTGCTTAAACAGTTGMAAAGACCCCTGATGGCGAGCGCTAATTCACAAAGCAATGATGG 5160
QY 5161 GAAACATGAATCGGTCTTAGGGAGGATCTGTCAAAAGTGGTCTGTGTTAAAAACAAGTGCC 5220
DB 5161 GAAACATGAATCGGTCTTAGGGAGGATCTGTCAAAAGTGGTCTGTGTTAAAAACAAGTGCC 5220
QY 5221 TCCTCCTCTCAGTGTCACTTTGATGTGTGCTTGAATTTCTCGGAAAACTGGGTGATGAG 5280
DB 5221 TCCTCCTCTCAGTGTCACTTTGATGTGTGCTTGAATTTCTCGGAAAACTGGGTGATGAG 5280
QY 5281 ACCCAGGATGAATTTGCCACAGGATTTGATGGAATCTTCCCTTCACTGCTCTTCAGCCA 5340
DB 5281 ACCCAGGATGAATTTGCCACAGGATTTGATGGAATCTTCCCTTCACTGCTCTTCAGCCA 5340
QY 5341 GTGCCAGTTTCTTTCTGATCATGTGATTTGAGTGAAGTCTGAGTCTGATATCAAAATC 5400
DB 5341 GTGCCAGTTTCTTTCTGATCATGTGATTTGAGTGAAGTCTGAGTCTGATATCAAAATC 5400
QY 5401 TTTAGAAATTTTGTAGTTTCTCGGACACAGGAAACCCAGCACTTAGCATATCAAAAT 5460
DB 5401 TTTAGAAATTTTGTAGTTTCTCGGACACAGGAAACCCAGCACTTAGCATATCAAAAT 5460
QY 5461 CTAATGTCTTAATGGCATATAAAGAGGCTTTTAAACACAGACTCCAGTTAGCTAAGTG 5520

DB 5461 CTAATGTCTTAATGGCATATAAAGAGGCTTTTAAACACAGACTCCAGTTAGCTAAGTG 5520
QY 5521 GTTCTGTAGTGGCGGTACTGTTGACGGGGCCCTGTGAGATGCCCCAGTTCCCTGAAAG 5580
DB 5521 GTTCTGTAGTGGCGGTACTGTTGACGGGGCCCTGTGAGATGCCCCAGTTCCCTGAAAG 5580
QY 5581 AAATGAAAAGGCCAGTTACCGGTAGGTGTGGGAAACATGGGCTAGATCATCAGGCAG 5640
DB 5581 AAATGAAAAGGCCAGTTACCGGTAGGTGTGGGAAACATGGGCTAGATCATCAGGCAG 5640
QY 5641 GACGAATGCTGGCTGTGGGTGGGAGCACCCAGCTTTGGCGTTTCTGTTCTTACC 5700
DB 5641 GACGAATGCTGGCTGTGGGTGGGAGCACCCAGCTTTGGCGTTTCTGTTCTTACC 5700
QY 5701 ACTGGTGTGTTTGTGACCAATTAATGAGTTGCTTAAACCTTTCTTTGCTACTATTTCCCTG 5760
DB 5701 ACTGGTGTGTTTGTGACCAATTAATGAGTTGCTTAAACCTTTCTTTGCTACTATTTCCCTG 5760
QY 5761 TTTGCAAAATGGTTTCACTTGAACCTTCTTCCACCTCCCAAGGACAAATTTCAACAGCCTA 5820
DB 5761 TTTGCAAAATGGTTTCACTTGAACCTTCTTCCACCTCCCAAGGACAAATTTCAACAGCCTA 5820
QY 5821 TTTGTAAAAAGATCACAGTCCCTTTAAAAAATATAAATCTGTAAGTCAAGAGTGATGCTTGA 5880
DB 5821 TTTGTAAAAAGATCACAGTCCCTTTAAAAAATATAAATCTGTAAGTCAAGAGTGATGCTTGA 5880
QY 5881 AAGAGCAGGAACCCAGGTAGATGTGAAATGTCAATCTCTTTGTTCTTAAAGAAAAAGCAAT 5940
DB 5881 AAGAGCAGGAACCCAGGTAGATGTGAAATGTCAATCTCTTTGTTCTTAAAGAAAAAGCAAT 5940
QY 5941 TCATAGCTTTTGGATATGACGCAACATACCAATAATCTTGACACATAGTTGGGAGTCGG 6000
DB 5941 TCATAGCTTTTGGATATGACGCAACATACCAATAATCTTGACACATAGTTGGGAGTCGG 6000
QY 6001 AAATTTGCAACAGCCAGTTTATAAACCCAGCTAGTTTGGGTATGATTTGTAAGAAAAAAA 6060
DB 6001 AAATTTGCAACAGCCAGTTTATAAACCCAGCTAGTTTGGGTATGATTTGTAAGAAAAAAA 6060
QY 6061 AGCTGGCCATTTCTGTAATTTGGGGAAATGATTTTCTTAACTTATATATCTTGTAGTCT 6120
DB 6061 AGCTGGCCATTTCTGTAATTTGGGGAAATGATTTTCTTAACTTATATATCTTGTAGTCT 6120
QY 6121 AGATTTATCATATGCTACTATCATCTGCTGCTTTTAAAGACTTAAAGAGATCAAGTAAT 6180
DB 6121 AGATTTATCATATGCTACTATCATCTGCTGCTTTTAAAGACTTAAAGAGATCAAGTAAT 6180
QY 6181 TTTTCTTTCTTTTGTAGACACTATATAGATCATCAAGGGTCTGCTCTTACAGGTGGAT 6240
DB 6181 TTTTCTTTCTTTTGTAGACACTATATAGATCATCAAGGGTCTGCTCTTACAGGTGGAT 6240
QY 6241 AGTGATATGATCTACAGTGAGGGGCAATTTTAAAACTTAAACATTCATGCTGTTTGG 6300
DB 6241 AGTGATATGATCTACAGTGAGGGGCAATTTTAAAACTTAAACATTCATGCTGTTTGG 6300
QY 6301 GGGTGGTATTTAAACGGCAGCACCTCTGATTTGCTTTTGGAGGGCTGGTGTGTTGAA 6360
DB 6301 GGGTGGTATTTAAACGGCAGCACCTCTGATTTGCTTTTGGAGGGCTGGTGTGTTGAA 6360
QY 6361 GTTCTGTCTCTCTCCAGTGGACTCTAACTTCTCTGATGACGCTGAGACACATTTGCTCT 6420
DB 6361 GTTCTGTCTCTCTCCAGTGGACTCTAACTTCTCTGATGACGCTGAGACACATTTGCTCT 6420
QY 6421 ATTGTCTGCAAAAATAAAGCCAAACACTGTCTCTGGGACAGGTTTTTCAATTTGTCAG 6480
DB 6421 ATTGTCTGCAAAAATAAAGCCAAACACTGTCTCTGGGACAGGTTTTTCAATTTGTCAG 6480
QY 6481 ATCTCTTTTGGCCCAATGAGTGTGTGACAAATACAGCTGCTTTCCAAAATTTGCTA 6540
DB 6481 ATCTCTTTTGGCCCAATGAGTGTGTGACAAATACAGCTGCTTTCCAAAATTTGCTA 6540
QY 6541 AATTTTGTAGCAGACTTTCTCTAGGTGCTTGCCTAAATGCCAGACTTTCTTTCTGTGAGAT 6600

6541	DB	AA	TTTTGACAGACTTTC	CTAGGTGCTTGCCCAATGTC	CCAGACTTTC	TTTCTGTTGTAAGAT	6600					
6601	QY	TA	AGTTGTGCTTGC	TGCCCTCTAGTGGTCAGTGT	TTAAATCCTTAACCTTAAACGGCTTAT	6660						
6601	DB	TA	AGTTGTGCTTGC	TGCCCTCTAGTGGTCAGTGT	TTAAATCCTTAACCTTAAACGGCTTAT	6660						
6661	QY	TTTT	CCCCCTGTTGGT	TGGGAAGTTGA	CGTTTGTAAATGGGCTCAATTTTCTAAATTAATTC	6720						
6661	DB	TTTT	CCCCCTGTTGGT	TGGGAAGTTGA	CGGTTTGTAAATGGGCTCAATTTTCTAAATTAATTC	6720						
6721	QY	TGA	AGAGATAA	TTTTTCCC	CCAGTAGTAGTATGTC	CCACCTTCAGTTTGC	CCAGATCC	TCGCC	6780			
6721	DB	TGA	AGAGATAA	TTTTTCCC	CCAGTAGTAGTATGTC	CCACCTTCAGTTTGC	CCAGATCC	TCGCC	6780			
6781	QY	TG	CTCAGAGACA	CTCAGAGAAC	CGGAAGCTG	CCCGGGCAATTCAGTCTAT	GAAATGATCTTT	6840				
6781	DB	TG	CTCAGAGACA	CTCAGAGAAC	CGGAAGCTG	CCCGGGCAATTCAGTCTAT	GAAATGATCTTT	6840				
6841	QY	CTT	GTGATTAA	GCGAAACG	AAGAACTG	GAATGTTTTAAATAGTGTACTCT	GTGTGTTACCCAGAA	6900				
6841	DB	CTT	GTGATTAA	GCGAAACG	AAGAACTG	GAATGTTTTAAATAGTGTACTCT	GTGTGTTACCCAGAA	6900				
6901	QY	AAAA	CAAAACAA	AAATCATG	TTATAAC	CTCTA	AAACAACTTCA	CAACAGCAAT	6960			
6901	DB	AAAA	CAAAACAA	AAATCATG	TTATAAC	CTCTA	AAACAACTTCA	CAACAGCAAT	6960			
6961	QY	TGG	TGTGTCTAG	CCGTTTTG	TTCTAAC	CCGATGTTATATA	AAAGAAATTTTTTTCATGCT	7020				
6961	DB	TGG	TGTGTCTAG	CCGTTTTG	TTCTAAC	CCGATGTTATATA	AAAGAAATTTTTTTCATGCT	7020				
7021	QY	TT	CAAAAAATG	TTTTATG	TCAGAAATATTTA	AGTCAGACATC	CTTATTCAGGTACTT	CAGC	7080			
7021	DB	TT	CAAAAAATG	TTTTATG	TCAGAAATATTTA	AGTCAGACATC	CTTATTCAGGTACTT	CAGC	7080			
7081	QY	TAC	CTTCTATATA	AAATATTTT	TGTTTTTCTTTT	TAAGATA	AAAAATGATGATGAGGAAAA	7140				
7081	DB	TAC	CTTCTATATA	AAATATTTT	TGTTTTTCTTTT	TAAGATA	AAAAATGATGATGAGGAAAA	7140				
7141	QY	AT	AAAA	CAATCTTAC	ATCTTCA	AGAGAGTGTAGTTAT	TGTGCAAGTTTTTCCGTTCAAAAT	7200				
7141	DB	AT	AAAA	CAATCTTAC	ATCTTCA	AGAGAGTGTAGTTAT	TGTGCAAGTTTTTCCGTTCAAAAT	7200				
7201	QY	TAT	TACCTCA	ATATTCAT	CTCAGAAC	GCATACAGGTAA	AGAACTTTATTTTTTAA	CCCA	7260			
7201	DB	TAT	TACCTCA	ATATTCAT	CTCAGAAC	GCATACAGGTAA	AGAACTTTATTTTTTAA	CCCA	7260			
7261	QY	TG	CAATTA	GTAAATATG	TAGTATCTAA	ATTTTTTTTGTGTTGT	TGTTAGATCTCTGC	7320				
7261	DB	TG	CAATTA	GTAAATATG	TAGTATCTAA	ATTTTTTTTGTGTTGT	TGTTAGATCTCTGC	7320				
7321	QY	CAG	ATCCTT	CGACTAG	CTTAAG	GATAAATATG	TAGCATGTGATG	TCAGTGGTATTTTT	7380			
7321	DB	CAG	ATCCTT	CGACTAG	CTTAAG	GATAAATATG	TAGCATGTGATG	TCAGTGGTATTTTT	7380			
7381	QY	ATT	CTTTT	TAGTGC	CAATTTA	ACTTG	ACCATTCTTATTTCTTTTCT	7440				
7381	DB	ATT	CTTTT	TAGTGC	CAATTTA	ACTTG	ACCATTCTTATTTCTTTTCT	7440				
7441	QY	TT	CTTTTTT	TGTTTTT	TGAGAC	CGAGTCTTG	CTGTCA	CCCTCGGCTGGAGTGAGTGGTG	7500			
7441	DB	TT	CTTTTTT	TGTTTTT	TGAGAC	CGAGTCTTG	CTGTCA	CCCTCGGCTGGAGTGAGTGGTG	7500			
7501	QY	CA	ATTT	CGGCTCA	CTG	CAGCCT	CCCTGTTCA	AGCAATATCTCTG	CCCTCAGCCT	7560		
7501	DB	CA	ATTT	CGGCTCA	CTG	CAGCCT	CCCTGTTCA	AGCAATATCTCTG	CCCTCAGCCT	7560		
7561	QY	CCC	CAGTAGT	TGGATTA	TCAG	GTACCTG	CCACAC	CCCGGCTAAATTTCTG	ATTTTTTTAG	7620		
7561	DB	CCC	CAGTAGT	TGGATTA	TCAG	GTACCTG	CCACAC	CCCGGCTAAATTTCTG	ATTTTTTTAG	7620		
7621	QY	TAG	AGATG	GGGTTT	CAC	ATGCTG	GGCAGG	CTGGTTTC	GAAATCTC	GATGATC	7680	
7621	DB	TAG	AGATG	GGGTTT	CAC	ATGCTG	GGCAGG	CTGGTTT	CGAACTC	CTC	GATGATC	7680

QY	7681	CGCTCACCTTGGCCTCCCATAGTGTTCGGCCTCCCATAGTCTGGAGTTACAGGCGTGAAGC	7740
Db	7681	CGCTCACCTTGGCCTCCCATAGTGTTCGGCCTCCCATAGTCTGGAGTTACAGGCGTGAAGC	7740
QY	7741	CACCGCGCCGGACAAAGTTCATTTGTTTATGACTTGCTATGCTTCCTGACTCTTTATC	7800
Db	7741	CACCGCGCCGGACAAAGTTCATTTGTTTATGACTTGCTATGCTTCCTGACTCTTTATC	7800
QY	7801	TTATTTAAAGCTACAGTATTTTAAATGCTGCACTCTTATGCTCTTATGATTTAGAAATGAA	7860
Db	7801	TTATTTAAAGCTACAGTATTTTAAATGCTGCACTCTTATGCTCTTATGATTTAGAAATGAA	7860
QY	7861	ATGAGAACTTATTTAGTAGTCTTGAGATGTGAAAGGAGCTATGACATCATGATGATAGGA	7920
Db	7861	ATGAGAACTTATTTAGTAGTCTTGAGATGTGAAAGGAGCTATGACATCATGATGATAGGA	7920
QY	7921	GGCTGCGTAGATTTGAAATTTTCACTCTTCCACTTACTCTGTCGACCCCTGGGCAAGT	7980
Db	7921	GGCTGCGTAGATTTGAAATTTTCACTCTTCCACTTACTCTGTCGACCCCTGGGCAAGT	7980
QY	7981	TATTTAAACCTTTTGTGCTTTTGTCTGCTGTTAAAGTGTGCTGCTTAAATTAATACATATTTTC	8040
Db	7981	TATTTAAACCTTTTGTGCTTTTGTCTGCTGTTAAAGTGTGCTGCTTAAATTAATACATATTTTC	8040
QY	8041	CCTAGGCGCTGTAGGAAGATTAATTAAGTTTGAAGTGTTCGCTGTTAAATTTTCTATTGAA	8100
Db	8041	CCTAGGCGCTGTAGGAAGATTAATTAAGTTTGAAGTGTTCGCTGTTAAATTTTCTATTGAA	8100
QY	8101	GATAGGCATTTCAATATTTCAAATATTCTATGTAAGTAAAGATGATAAAGAACTGATGAGAA	8160
Db	8101	GATAGGCATTTCAATATTTCAAATATTCTATGTAAGTAAAGATGATAAAGAACTGATGAGAA	8160
QY	8161	ATCCTATGTGATAGTAGATCGAGAAAGCAAAAGAGGAAAGACCTGTTTCTTAAATAA	8220
Db	8161	ATCCTATGTGATAGTAGATCGAGAAAGCAAAAGAGGAAAGACCTGTTTCTTAAATAA	8220
QY	8221	ATAGATATTTGATCTATTTTCAGTGTCTTTTCATACACTTCTATAATAAGTGCCATTTCTT	8280
Db	8221	ATAGATATTTGATCTATTTTCAGTGTCTTTTCATACACTTCTATAATAAGTGCCATTTCTT	8280
QY	8281	GCCTTAGGTGAAAAACCATCAAAATGTGAATTTTGTGAATATGCTCGAGCCCAAGAACA	8340
Db	8281	GCCTTAGGTGAAAAACCATCAAAATGTGAATTTTGTGAATATGCTCGAGCCCAAGAACA	8340
QY	8341	TCTCTGAGGTATCACTTTGGAGAGACATCAAGAGGAAAAACAACCGATGCTGCTGCTGAA	8400
Db	8341	TCTCTGAGGTATCACTTTGGAGAGACATCAAGAGGAAAAACAACCGATGCTGCTGCTGAA	8400
QY	8401	GTCAAGAACGATGGTAAAAATCAGGACACTGGAAGTGCATATTAAACCGCTGACAGTGGC	8460
Db	8401	GTCAAGAACGATGGTAAAAATCAGGACACTGGAAGTGCATATTAAACCGCTGACAGTGGC	8460
QY	8461	CAAAACAAAAATTTGAAAAAGATTTTTTGTGATGCTGTCGCAAGATGTTACAGGAGTCCACCT	8520
Db	8461	CAAAACAAAAATTTGAAAAAGATTTTTTGTGATGCTGTCGCAAGATGTTACAGGAGTCCACCT	8520
QY	8521	GCAAGCAGCTTAAGGAGATGCCTCTGTTTTCAGAAATGTTCTGGGACCGCTGCTC	8580
Db	8521	GCAAGCAGCTTAAGGAGATGCCTCTCTGTTTTCAGAAATGTTCTGGGACCGCTGCTC	8580
QY	8581	TCACCAGCACAAAGATATCTCAGGATTTTCCATATAAAATGCACTGATGACAGTGTGAT	8640
Db	8581	TCACCAGCACAAAGATATCTCAGGATTTTCCATATAAAATGCACTGATGACAGTGTGAT	8640
QY	8641	AAAGTGAATAAAAACCTTACCTCGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTT	8700
Db	8641	AAAGTGAATAAAAACCTTACCTCGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTT	8700
QY	8701	GAAACTCAGGCAATAAATCTCTGTAGAACCAAGGCGGATGTTTACTCTCTCGGAT	8760
Db	8701	GAAACTCAGGCAATAAATCTCTGTAGAACCAAGGCGGATGTTTACTCTCTCGGAT	8760

QY 8761 GGCAGTACCACCATAACTTGAAGTTAGTCCCAAGAGAACGAGAGCCGAGCT 8820
DB 8761 GGCAGTACCACCATAACTTGAAGTTAGTCCCAAGAGAACGAGAGCCGAGCT 8820
QY 8821 GACTCGAGATACAGGCCAAGTGTGGATGTACGAAACACCTTTAAATTTATCCGTGGGG 8880
DB 8821 GACTCGAGATACAGGCCAAGTGTGGATGTACGAAACACCTTTAAATTTATCCGTGGGG 8880
QY 8881 GCTCTTCAAAATGCGCGCAATTTCTTTGAGTAAAGTTTGATTCGAAGTATCACTGT 8940
DB 8881 GCTCTTCAAAATGCGCGCAATTTCTTTGAGTAAAGTTTGATTCGAAGTATCACTGT 8940
QY 8941 CCAATTTGTACCTTCAAGACATTTTATCCAGAGTTTAAATGATGACACGAGACTGGAG 9000
DB 8941 CCAATTTGTACCTTCAAGACATTTTATCCAGAGTTTAAATGATGACACGAGACTGGAG 9000
QY 9001 CATAAATACAATCTGACGTTTCATAAAACCTGCGAAACAGTCTTGTGTAGAAAGTCA 9060
DB 9001 CATAAATACAATCTGACGTTTCATAAAACCTGCGAAACAGTCTTGTGTAGAAAGTCA 9060
QY 9061 CGTACCGGATGCCCGCAGCTGTGTGGGAAAGATGTGCCTCCCTCCCTAGTTTCTGT 9120
DB 9061 CGTACCGGATGCCCGCAGCTGTGTGGGAAAGATGTGCCTCCCTCCCTAGTTTCTGT 9120
QY 9121 AAACCCAGCCCAAGTCTGTTCCCGCGCAGTCCAAATCCTGCCATCTCGGAAGGG 9180
DB 9121 AAACCCAGCCCAAGTCTGTTCCCGCGCAGTCCAAATCCTGCCATCTCGGAAGGG 9180
QY 9181 AAGCAGAGCCCTCTGGGCGAGCAAGGCCCTCTGACTTCAGGGATAGACTCTAGCACT 9240
DB 9181 AAGCAGAGCCCTCTGGGCGAGCAAGGCCCTCTGACTTCAGGGATAGACTCTAGCACT 9240
QY 9241 TTAGCCCCAAGTAACTGAAGTCCCAAGTCCCAAGTCCCAAGTCCCAAGTCCCAAGTCC 9300
DB 9241 TTAGCCCCAAGTAACTGAAGTCCCAAGTCCCAAGTCCCAAGTCCCAAGTCCCAAGTCC 9300
QY 9301 GCACCCGACACAGCATCTCAGATGTTTCTTAAACAGTGTTCCTCCCTGACCGAT 9360
DB 9301 GCACCCGACACAGCATCTCAGATGTTTCTTAAACAGTGTTCCTCCCTGACCGAT 9360
QY 9361 AAGACAAAAGACCCGAGACAAAATTTGAACCTCTTCCAGTAGCTCTCTCAGGCCACC 9420
DB 9361 AAGACAAAAGACCCGAGACAAAATTTGAACCTCTTCCAGTAGCTCTCTCAGGCCACC 9420
QY 9421 CTCGGCAGCAGTAAACATCAATGTTTCCATCGACTACCCCGCAGAACGACAGCCCGTGG 9480
DB 9421 CTCGGCAGCAGTAAACATCAATGTTTCCATCGACTACCCCGCAGAACGACAGCCCGTGG 9480
QY 9481 GCACCTCCGGGAGAGACTATTTCTGTAATCGAGTGCAGCAATATCTCAGCAGAAATTT 9540
DB 9481 GCACCTCCGGGAGAGACTATTTCTGTAATCGAGTGCAGCAATATCTCAGCAGAAATTT 9540
QY 9541 GGTGAGCCCCCTTCCAAAAGACTGAAGTCCAGCGTGTGTCCTTGACCTTGACCCAGCC 9600
DB 9541 GGTGAGCCCCCTTCCAAAAGACTGAAGTCCAGCGTGTGTCCTTGACCTTGACCCAGCC 9600
QY 9601 GGGGCCAAATACAGAGAGGCTATGACCTTCCCAAGTACATATGTTGTCAGAGCATCA 9660
DB 9601 GGGGCCAAATACAGAGAGGCTATGACCTTCCCAAGTACATATGTTGTCAGAGCATCA 9660
QY 9661 TCACCTGTTACCGCAGGACTGTGTATCCGTGCGAGGCGCTGCCTCCCAACCAAGGTTTC 9720
DB 9661 TCACCTGTTACCGCAGGACTGTGTATCCGTGCGAGGCGCTGCCTCCCAACCAAGGTTTC 9720
QY 9721 CTGAGCTCCAGGAGGTGATTTCTCCAAATGTGCTGACTGTTCAGAGCCCTATGTTGGC 9780
DB 9721 CTGAGCTCCAGGAGGTGATTTCTCCAAATGTGCTGACTGTTCAGAGCCCTATGTTGGC 9780
QY 9781 TCGGGCCACTTTACACTTGTGTGCTGCTGCTAGTCCAGCATCCAGCTCGAGCTTAGAA 9840
DB 9781 TCGGGCCACTTTACACTTGTGTGCTGCTGCTAGTCCAGCATCCAGCTCGAGCTTAGAA 9840
QY 9841 GGTATTTGCATGAGGGCGTGTGTTTAAATGGCTGCTACAGTGAATTAATAGTAATCCA 9900

DB 9841 GGTATTTGCATGAGGGCGTGTGTTTAAATGGCTGCTACAGTGAATTAATAGTAATCCA 9900
QY 9901 GGCATTTCTCAGTGGAGATGGTACCACTCCCAAGGTTGGGGTAGGCGCCAGAAATTTCT 9960
DB 9901 GGCATTTCTCAGTGGAGATGGTACCACTCCCAAGGTTGGGGTAGGCGCCAGAAATTTCT 9960
QY 9961 TGGGGTTCACAGAGAGAGCAATTTCTTAGATACGGCAGTGG 10000
DB 9961 TGGGGTTCACAGAGAGAGCAATTTCTTAGATACGGCAGTGG 10000

RESULT 2
AF312915/c 121143 bp DNA linear PRI 13-JUN-2001
LOCUS Homo sapiens chromosome 20 clones 97 and 127, complete sequence.
DEFINITION AF312915
ACCESSION AF312915
VERSION AF312915.1 GI:11094030
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 121143)
AUTHORS Collins, C., Volik, S., Kowbel, D., Ginzinger, D., Ylstra, B.,
Cloutier, T., Hawkins, T., Predki, P., Martin, C., Wernick, M.,
Kuo, W.B., Alberts, A. and Gray, J.W.
Comprehensive genome sequence analysis of a breast cancer amplicon
Genome Res. 11 (6), 1034-1042 (2001)
MEDLINE 21275464
PUBMED 11381030
REFERENCE 2 (bases 1 to 121143)
AUTHORS Volik, S., Collins, C., Gray, J., Wernick, M., Kowbel, D., Stultz, K. and
Martin, C.
Direct Submission
TITLE Submitted (10-OCT-2000) Cancer Genetics, UCSF Cancer Center, 2340
JOURNAL Sutter St., Rm. S151, San Francisco, CA 94706, USA
FEATURES
source
1..121143
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20q13.2"
/clone="97; 127; LBNL hl17 and h79"
/note="Isolated in Collins lab and sequenced by the Genome
Center at LBNL"
complement(1..66093)
/note="overlaps Homo sapiens chromosome 20 clone
RP5-823G15 deposited in GenBank Accession Number
AU354993.19"
194..468
repeat_region
/rpt_family="Alu"
/rpt_type=dispersed
complement(513..653)
repeat_region
/rpt_family="MER5"
/rpt_type=dispersed
1053..1153
repeat_region
/rpt_family="MER20"
/rpt_type=dispersed
1254..1330
repeat_region
/rpt_family="Alu"
/rpt_type=dispersed
complement(1356..1676)
repeat_region
/rpt_family="Alu"
/rpt_type=dispersed
1681..1866
repeat_region
/rpt_family="Alu"
/rpt_type=dispersed
2485..5421
misc_feature
/note="similar to Homo sapiens chromosome 5 clone
CTC-269F1 deposited in GenBank Accession Number
AC011341.2"

```

repeat_region 2594..2879
/rpt_family="Alu"
/rpt_type=dispersed
misc_feature complement(2896..4437)
/notes="similar to flow-sorted chromosome 20 HindIII
fragment, SC20pF17C6 deposited in GenBank Accession Number
294682.1"
repeat_region 3106..3412
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 3889..4206
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 4449..4764
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(5671..5987)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 6723..7004
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(7486..7625)
/rpt_family="L1"
/rpt_type=dispersed
repeat_region 7716..7965
/rpt_family="Alu"
/rpt_type=dispersed
misc_feature complement(8590..118606)
/notes="similar to Mus musculus clone CT7-510C4 deposited
in GenBank Accession Number AC023610.3"
misc_feature complement(9852..10109)
/notes="similar to STS WI-31010 deposited in GenBank
Accession Number G24374.1"
mRNA <10244..>10904
/product="hypothetical"
/notes="similar to Homo sapiens hypothetical protein
OVC10-2 deposited in GenBank Accession Numbers AF230201.1,
AP00365 and in NM_018692.1"
repeat_region /evidence=not_experimental
15301..15588
/rpt_family="Alu"
/rpt_type=dispersed
misc_feature 16873..17073
/notes="similar to Homo sapiens STS WI-16697 deposited in
GenBank Accession Number G21490.1"
repeat_region 17577..17857
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(19034..19333)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 20136..20719
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 20776..21071
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(22344..22638)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 22726..22794
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(22984..23268)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(23349..23646)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(24008..24189)
/rpt_family="MIR"
/rpt_type=dispersed

```

Query Match 100.0%; Score 10000; DB 9; Length 121143;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 10000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db	105595	AGCACTGATTTAAACTCGATTTTAACTGGATGAAATTTCTGATTTAATAAGTACTGACT	105536
Qy	2221	GGATAAAATGCCAATGATTTAAATTAACAAGCACGCTTTTAAACAGGATGCCCTATATATTAGT	2280
Db	105535	GGATAAAATGCCAATGATTTAAATTAACAAGCACGCTTTTAAACAGGATGCCCTATATATTAGT	105476
Qy	2281	TAAAGTGAAGCAATTGAATTTAGGTACCTTTCTGCTGCGTGGAAAGACCGTATGACTC	2340
Db	105475	TAAAGTGAAGCAATTGAATTTAGGTACCTTTCTGCTGCGTGGAAAGACCGTATGACTC	105416
Qy	2341	ACCCACACCAAGCCTTCTCTGCTGCTGAGTGAACCGTTTCTGTTTCTCTCT	2400
Db	105415	ACCCACACCAAGCCTTCTCTGCTGCTGAGTGAACCGTTTCTGTTTCTCTCT	105356
Qy	2401	AGGGTTTGGAAATCCCTTGTCTCCAGGTTCCTGGGATTGACTTCTTGCTCAATTTGAACA	2460
Db	105355	AGGGTTTGGAAATCCCTTGTCTCCAGGTTCCTGGGATTGACTTCTTGCTCAATTTGAACA	105296
Qy	2461	CTCATTCATTTGGAGACAAGAGAACTAATGCTTTGCTGCTGATTCATATTTGAATCGAGGC	2520
Db	105295	CTCATTCATTTGGAGACAAGAGAACTAATGCTTTGCTGCTGATTCATATTTGAATCGAGGC	105236
Qy	2521	ATTGGGAACCTCTATGCTGCTTGTGTTGTTGGAAGAACCAAGTGACACCATCACTGAGCTTC	2580
Db	105235	ATTGGGAACCTCTATGCTGCTTGTGTTGTTGGAAGAACCAAGTGACACCATCACTGAGCTTC	105176
Qy	2581	TAAAGTTTCAAGAAAGTTAGAGGACTATACACTTTCTTTTGAACCTTTTATATAAATATT	2640
Db	105175	TAAAGTTTCAAGAAAGTTAGAGGACTATACACTTTCTTTTGAACCTTTTATATAAATATT	105116
Qy	2641	TGCTCTGGTTTTTGGAAACCCAGGGCTGTTAGAGGGGTGAGTGACAAAGTCTTCAAGTGGC	2700
Db	105115	TGCTCTGGTTTTTGGAAACCCAGGGCTGTTAGAGGGGTGAGTGACAAAGTCTTCAAGTGGC	105056
Qy	2701	CTTATTTCCAACTCCAGAAATGCCCCAACGAACTTTGAGATTATATGCAATCGAAAGTGA	2760
Db	105055	CTTATTTCCAACTCCAGAAATGCCCCAACGAACTTTGAGATTATATGCAATCGAAAGTGA	104996
Qy	2761	CAGGAAACATGCCAACTCAATCCCTCTTAATGATGATGAGTGGCCAGAGTGTATGGCA	2820
Db	104995	CAGGAAACATGCCAACTCAATCCCTCTTAATGATGATGAGTGGCCAGAGTGTATGGCA	104936
Qy	2821	GCTCTCTTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATGAAAGGACCGCTGTTG	2880
Db	104935	GCTCTCTTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATGAAAGGACCGCTGTTG	104876
Qy	2881	TTCCATTTCCAGGTACACAAGAAATAATGTCTATCCAAATCGAGGGGTATATGCCCTTGG	2940
Db	104875	TTCCATTTCCAGGTACACAAGAAATAATGTCTATCCAAATCGAGGGGTATATGCCCTTGG	104816
Qy	2941	ATTGCAATGTTCTGACGACAGACCTTCCACATTCAGAGACCTTAAATAACATGCTTAA	3000
Db	104815	ATTGCAATGTTCTGACGACAGACCTTCCACATTCAGAGACCTTAAATAACATGCTTAA	104756
Qy	3001	TGCAACACCGGGCTACCTCTGTAACACGAGTCTTCTCGGTTTGAAGCAGAGTATCTCA	3060
Db	104755	TGCAACACCGGGCTACCTCTGTAACACGAGTCTTCTCGGTTTGAAGCAGAGTATCTCA	104696
Qy	3061	GTCCGCTTATATAAAGTCAAGTCCGAAACAGACCTCCCAAGGAAAGAAATTCAGAGAAA	3120
Db	104695	GTCCGCTTATATAAAGTCAAGTCCGAAACAGACCTCCCAAGGAAAGAAATTCAGAGAAA	104636
Qy	3121	ATGAATTTAGTGTGAGGTATGTGGGCAGACATTTTAGAGTTCGCTTTTGTATGTTGAGATCC	3180
Db	104635	ATGAATTTAGTGTGAGGTATGTGGGCAGACATTTTAGAGTTCGCTTTTGTATGTTGAGATCC	104576
Qy	3181	ACATGAGAAACACAAGAGTTCTTTTCACTTACGGGTGTAAACATGTCCGGAAGAAATTCA	3240
Db	104575	ACATGAGAAACACAAGAGTTCTTTTCACTTACGGGTGTAAACATGTCCGGAAGAAATTCA	104516
Qy	3241	AGAGCCTTGGTTTCTTAAAAATCACAATCGGACACATATGCGAAATCGGGGGCCAGAA	3300
Db	104515	AGAGCCTTGGTTTCTTAAAAATCACAATCGGACACATATGCGAAATCGGGGGCCAGAA	104456
Qy	3301	GCAAACTGCAGCAAGGCTTTCGAGAGTAGTCCAGCAACGATCAACGAGGTGCTGCCAGGTGC	3360
Db	104455	GCAAACTGCAGCAAGGCTTTCGAGAGTAGTCCAGCAACGATCAACGAGGTGCTGCCAGGTGC	104396
Qy	3361	ACCGGCGGAGAGCATCTCCTCTCTTACAAAATCTGCATGTTGTGGCTTCTTATTTTC	3420
Db	104395	ACCGGCGGAGAGCATCTCCTCTCTTACAAAATCTGCATGTTGTGGCTTCTTATTTTC	104336
Qy	3421	CAAAATGAAGAAAGTCTAAATTTAGACACCGCAAGGTGCACACAAAATACTGCTTTCGGTA	3480
Db	104335	CAAAATGAAGAAAGTCTAAATTTAGACACCGCAAGGTGCACACAAAATACTGCTTTCGGTA	104276
Qy	3481	CCAGCAGCGCGCAGACAGACTCTCCACAGAGGAAATGCGCTCTCCAGGGAGGACTTCC	3540
Db	104275	CCAGCAGCGCGCAGACAGACTCTCCACAGAGGAAATGCGCTCTCCAGGGAGGACTTCC	104216
Qy	3541	TGCAGTTGTTCAACTTTGAGACCAAAATCTCACCTGAAACGGGGAAGAACCTGTTCAGAT	3600
Db	104215	TGCAGTTGTTCAACTTTGAGACCAAAATCTCACCTGAAACGGGGAAGAACCTGTTCAGAT	104156
Qy	3601	GCATCCCTCAGCTCGATCCGTTTCCACCACTTCCAGGCTTGGCAGCTGGCTACCAAGGAA	3660
Db	104155	GCATCCCTCAGCTCGATCCGTTTCCACCACTTCCAGGCTTGGCAGCTGGCTACCAAGGAA	104096
Qy	3661	AGTTTCCCAATTTGCCCAAGAAAGTGAAGGAAATCGGGGCAAGAGGAGCACCAGCAACGACG	3720
Db	104095	AGTTTCCCAATTTGCCCAAGAAAGTGAAGGAAATCGGGGCAAGAGGAGCACCAGCAACGACG	104036
Qy	3721	ATTTCGAGTTCCGAGAAAGGAGCTTTGGAGAAAATAATTAAGGACAGTTGTGACAGGCTCTCGC	3780
Db	104035	ATTTCGAGTTCCGAGAAAGGAGCTTTGGAGAAAATAATTAAGGACAGTTGTGACAGGCTCTCGC	103976
Qy	3781	RAGAGAAAGAGAGTCCAAACACTCCCAACGCGAAGCGGCTCCGTTGGAGCGCGATCCCA	3840
Db	103975	RAGAGAAAGAGAGTCCAAACACTCCCAACGCGAAGCGGCTCCGTTGGAGCGCGATCCCA	103916
Qy	3841	AGTTTACCCAGTAGAGAGGAGAGCCCACTCAGTCTCCAGAGTCCGCGCAAAAGCTTTTCAGAA	3900
Db	103915	AGTTTACCCAGTAGAGAGGAGAGCCCACTCAGTCTCCAGAGTCCGCGCAAAAGCTTTTCAGAA	103856
Qy	3901	CCTACCAACAGCTGGTCTTTTCACATCCAGGTTCCAAAGAGGACCGGAGGGCGGCGCGG	3960
Db	103855	CCTACCAACAGCTGGTCTTTTCACATCCAGGTTCCAAAGAGGACCGGAGGGCGGCGCGG	103796
Qy	3961	AGTCCGCCACCATGTTCTGTGAGAGGAGCGCGGGGAGCTGTTCTCTGACCTCGCCG	4020
Db	103795	AGTCCGCCACCATGTTCTGTGAGAGGAGCGCGGGGAGCTGTTCTCTGACCTCGCCG	103736
Qy	4021	CCCCTCTGATGAAAATGAGCGCGTGCATCGAGGGGAAGTGGTTCTGAAAGCGGATCTG	4080
Db	103735	CCCCTCTGATGAAAATGAGCGCGTGCATCGAGGGGAAGTGGTTCTGAAAGCGGATCTG	103676
Qy	4081	AGGATGGGTTCCCGAAGAAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCTCGTGTGT	4140
Db	103675	AGGATGGGTTCCCGAAGAAATCCATCTGGGTAAGCTGCCCTGTCTCCGTCCTCGTGTGT	103616
Qy	4141	TCCGCTGTGTCTGTCTGTCTCCCGTCTCCCGCTCTCTATTTCCCATCTCCAGACACCG	4200
Db	103615	TCCGCTGTGTCTGTCTGTCTCCCGTCTCCCGCTCTCTATTTCCCATCTCCAGACACCG	103556
Qy	4201	TGGCCAGGAATGGGTTTTCGAGAGCCAGAGTCAAGTCCAGGCTCTTTTGGTATCACTCT	4260
Db	103555	TGGCCAGGAATGGGTTTTCGAGAGCCAGAGTCAAGTCCAGGCTCTTTTGGTATCACTCT	103496
Qy	4261	GTGTAAGTCAATTAACCTCTCAGGGCTTAAATTTTCTCAATTTCTGTAATAACAGGGTTGA	4320
Db	103495	GTGTAAGTCAATTAACCTCTCAGGGCTTAAATTTTCTCAATTTCTGTAATAACAGGGTTGA	103436
Qy	4321	GTTAAGAGTCTCTTGTCTGAAATAATATATATATTTTAAACGCTGTATCGTTTTGC	4380
Db	103435	GTTAAGAGTCTCTTGTCTGAAATAATATATATATTTTAAACGCTGTATCGTTTTGC	103376

QY	4381	TCACAAAACACACTTTTAAAAAATAAATCTGTGCATCCAGCCCAAAATGCACCTGCTCT	4440	Db	102295	CTAATGCTTAATGSCATCATAAAAAGAGGCTTTAAAAACACAGACTCCAGTTAGCTAAGTG	102236
Db	103375	TCACAAAACACACTTTTAAAAAATAAATCTGTGCATCCAGCCCAAAATGCACCTGCTCT	103316				
QY	4441	TAACTGGGGCGATTTTGTTCCTCAATCAGTATCTGGCAATGTCTGGAGGCAATTTGGTTGT	4500	QY	5521	GTTTCTGCTAGTCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCTCTGAAG	5580
Db	103315	TAACTGGGGCGATTTTGTTCCTCAATCAGTATCTGGCAATGTCTGGAGGCAATTTGGTTGT	103256	Db	102235	GTTTCTGCTAGTCCGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCTCTGAAG	102176
QY	4501	CATACCTGTGTGTGGGTGTGCTGCTGGCATCCAGTGGGAGAGGCCAGGGACACTGCT	4560	QY	5581	AAATGAAAGGCCAGTTTACCGGTAGGTGTGGGAAACATCGGGCTAGATCATCAGGCGAG	5640
Db	103255	CATACCTGTGTGTGGGTGTGCTGCTGGCATCCAGTGGGAGAGGCCAGGGACACTGCT	103196	Db	102175	AAATGAAAGGCCAGTTTACCGGTAGGTGTGGGAAACATCGGGCTAGATCATCAGGCGAG	102116
QY	4561	CAGCATGTTACAGTCACAGGACAGCCCATCATCAAGAATAATCTGCTCCCAATGTC	4620	QY	5641	GACAGAAATGCTGCTGTGGGTGGGAGCAACCCAGCTTGGCGTTGAGTTCCTGTTCTACC	5700
Db	103195	CAGCATGTTACAGTCACAGGACAGCCCATCATCAAGAATAATCTGCTCCCAATGTC	103136	Db	102115	GACAGAAATGCTGCTGTGGGTGGGAGCAACCCAGCTTGGCGTTGAGTTCCTGTTCTACC	102056
QY	4621	AATAGTTTGGAGCAATTTAGAGAGCCCTAGCCCTTCACTTAAAGTTTCTGGCGTTCTGATCT	4680	QY	5701	ACTGCTGTTTGTGTCACAAATATGAGTGTGCTTAAACCTTCTTTGCTACTATTTCCCTG	5760
Db	103135	AATAGTTTGGAGCAATTTAGAGAGCCCTAGCCCTTCACTTAAAGTTTCTGGCGTTCTGATCT	103076	Db	102055	ACTGCTGTTTGTGTCACAAATATGAGTGTGCTTAAACCTTCTTTGCTACTATTTCCCTG	101996
QY	4681	TTTTCTGTAGTGAATTTCTAGTGGCCATAAAGGTACTGGGAGTGATCAATAGAGCCAG	4740	QY	5761	TTTGCAGAAATGGTTTCAATTTGACCCCTGCTTCCACCTCCCAAGGACAAATTTCAACGCCTA	5820
Db	103075	TTTTCTGTAGTGAATTTCTAGTGGCCATAAAGGTACTGGGAGTGATCAATAGAGCCAG	103016	Db	101995	TTTGCAGAAATGGTTTCAATTTGACCCCTGCTTCCACCTCCCAAGGACAAATTTCAACGCCTA	101936
QY	4741	GAATATTTTGGGAGCGTTTGGTGTCTCCAAAACCTTCTGCTCTGCTGCGCAAG	4800	QY	5821	TTTGTAAAAAGATCACAGTCCCTTTAAAAAATATAAATCTGTAAAGTCAGAGGTGATGCTTGA	5880
Db	103015	GAATATTTTGGGAGCGTTTGGTGTCTCCAAAACCTTCTGCTCTGCTGCGCAAG	102956	Db	101935	TTTGTAAAAAGATCACAGTCCCTTTAAAAAATATAAATCTGTAAAGTCAGAGGTGATGCTTGA	101876
QY	4801	CTAGTATCCATTTATAGGTACTCCAGGAACCCAAATGATTTGTCATAAAATAAAGGAAT	4860	QY	5881	AAGACAGGAACCCAGGTAGATGTGAAATGTCTCTTTGTTCTTAAAGAAAGGCATT	5940
Db	102955	CTAGTATCCATTTATAGGTACTCCAGGAACCCAAATGATTTGTCATAAAATAAAGGAAT	102896	Db	101875	AAGACAGGAACCCAGGTAGATGTGAAATGTCTCTTTGTTCTTAAAGAAAGGCATT	101816
QY	4861	GTGAGCACACTGAAGACATTTTAAAGGCTCATTTGCTCAGAGAAATTTTCACTGTAC	4920	QY	5941	TCATAGCTTTTGGATATGACGCAACATACCAATAATCTGACACATAGTTGGGAGTCGG	6000
Db	102895	GTGAGCACACTGAAGACATTTTAAAGAGGCTCATTTGCTCAGAGAAATTTTCACTGTAC	102836	Db	101815	TCATAGCTTTTGGATATGACGCAACATACCAATAATCTGACACATAGTTGGGAGTCGG	101756
QY	4921	TAGTGGCATTTATAGAAAGAGAGTGATCACTGAAGGCATCTCACATAATATTCTTGA	4980	QY	6001	AAATGTCAACAAACGCCAGTTTATAAACCCAGCTAGTTTGGGTATGATTTGTAAGAAAAAA	6060
Db	102835	TAGTGGCATTTATAGAAAGAGAGTGATCACTGAAGGCATCTCACATAATATTCTTGA	102776	Db	101755	AAATGTCAACAAACGCCAGTTTATAAACCCAGCTAGTTTGGGTATGATTTGTAAGAAAAAA	101696
QY	4981	GCCCTGGTGGGCTTATCTAGGGCAAGGATTTCCACCTGTGTGTTGGAGTTGGCCCATCC	5040	QY	6061	AGCTGGCCATTTCTGTATTTGGGAAATTTGATTTTCTTAACTTATATATCTTAGTAGTCT	6120
Db	102775	GCCCTGGTGGGCTTATCTAGGGCAAGGATTTCCACCTGTGTGTTGGAGTTGGCCCATCC	102716	Db	101695	AGCTGGCCATTTCTGTATTTGGGAAATTTGATTTTCTTAACTTATATATCTTAGTAGTCT	101636
QY	5041	TCACCTGTAGCCAGAGCTTCTCTATCAGAGTTTAGTATTTTGTGTAATAGAGATCTTG	5100	QY	6121	AGATTTATCATATCTGATCATCTGCTCTTTTAAAGCTTAAAGAGATCAAGTAAAT	6180
Db	102715	TCACCTGTAGCCAGAGCTTCTCTATCAGAGTTTAGTATTTTGTGTAATAGAGATCTTG	102656	Db	101635	AGATTTATCATATCTGATCATCTGCTCTTTTAAAGCTTAAAGAGATCAAGTAAAT	101576
QY	5101	CTGCTTAAACAGTTGAAAGACCCCTGATGGGAGCCGTAATTCACAGCGCAATGATGG	5160	QY	6181	TTTTTTTTTCTTTTGTAGACACTATATAGATCATCAAGGGTGTCTGCTTTACAGGTGGAT	6240
Db	102655	CTGCTTAAACAGTTGAAAGACCCCTGATGGGAGCCGTAATTCACAGCGCAATGATGG	102596	Db	101575	TTTTTTTTTCTTTTGTAGACACTATATAGATCATCAAGGGTGTCTGCTTTACAGGTGGAT	101516
QY	5161	GAACATGAATCCGTCTTAGGGAGGATCTGTCAAAGTGGTCTTGGTTAAAAACAAGTGCC	5220	QY	6241	AGTGATATGATCTACAGTGAGGGGACATTTTAAAACTTAAACATTTCACTGTTTGG	6300
Db	102595	GAACATGAATCCGTCTTAGGGAGGATCTGTCAAAGTGGTCTTGGTTAAAAACAAGTGCC	102536	Db	101515	AGTGATATGATCTACAGTGAGGGGACATTTTAAAACTTAAACATTTCACTGTTTGG	101456
QY	5221	TCCCTCTCAGTGTCATTTGTTGCTGCTTGAATTTCTTCGGAAAACTGGGTGTATGAG	5280	QY	6301	GGTGGTATTTTAAACGGCAGCACCTCTGATTTGCTTTTGGAGGGCTGTTGTTGTA	6360
Db	102535	TCCCTCTCAGTGTCATTTGTTGCTGCTTGAATTTCTTCGGAAAACTGGGTGTATGAG	102476	Db	101455	GGTGGTATTTTAAACGGCAGCACCTCTGATTTGCTTTTGGAGGGCTGTTGTTGTA	101396
QY	5281	ACCCAGGATGAATTTGCCACACAGATTTGATTTGAGTCTTCTTCCCTGCTCTCAGCCA	5340	QY	6361	GTTCTGTCTCTCTCCAGTGAGCTCTAACTTCTCTGATGACGTCGAGACACATTTGCTCT	6420
Db	102475	ACCCAGGATGAATTTGCCACACAGATTTGATTTGAGTCTTCTTCCCTGCTCTCAGCCA	102416	Db	101395	GTTCTGTCTCTCTCCAGTGAGCTCTAACTTCTCTGATGACGTCGAGACACATTTGCTCT	101336
QY	5341	GTGCCAGTTTCTTTCTGATCATGTGATGAGCTGAGAACTGTAGTCTGTATATCAAAATC	5400	QY	6421	ATTGCTCTGCAGAAACTTAAAGCCAAACACTGTCTCTGAGGAGAGGTTTCAATTTGTCAG	6480
Db	102415	GTGCCAGTTTCTTTCTGATCATGTGATGAGCTGAGAACTGTAGTCTGTATATCAAAATC	102356	Db	101335	ATTGCTCTGCAGAAACTTAAAGCCAAACACTGTCTCTGAGGAGAGGTTTCAATTTGTCAG	101276
QY	5401	TTTAGAATGTTTGTAGTTTCTTGGGACACAGGAAACCCAGCACTTAGCATACTACAAT	5460	QY	6481	ATCTCTTTCGCCCACTGAGTGTGTTGGACAAATACAGGCTGCTTTCCAAAACTTTGCTA	6540
Db	102355	TTTAGAATGTTTGTAGTTTCTTGGGACACAGGAAACCCAGCACTTAGCATACTACAAT	102296	Db	101275	ATCTCTTTCGCCCACTGAGTGTGTTGGACAAATACAGGCTGCTTTCCAAAACTTTGCTA	101216
QY	5461	CTAATGTCTTAATGGCATATAAAAAAGAGGCTTTTAAACACAGACTCCAGTTAGCTAAGTG	5520	QY	6541	AAATTTTGACAGACTTTTCTCTAGGTGCTTGCCTAATGCCAGACTTTCTTTCTGTGAAGAT	6600

QY	8761	GGCAGTACCACCATTAACCTTGAAGTTAGTCCCAAGAGAAAGCAACCGAGAGCCGAGCT	8820
Db	98995	GGCAGTACCACCATTAACCTTGAAGTTAGTCCCAAGAGAAAGCAACCGAGAGCCGAGCT	98936
QY	8821	GACTGCAGATACAGGCCAAGTGTGGAATCTGACGAAAAACCTTTAAATTTATCCGTGGGG	8880
Db	98935	GACTGCAGATACAGGCCAAGTGTGGAATCTGACGAAAAACCTTTAAATTTATCCGTGGGG	98876
QY	8881	GCTCTTCAAAATGCGCGCAATTTCTTTGAGTAAAGTTTGATTTCCAGTAGTACCTGT	8940
Db	98875	GCTCTTCAAAATGCGCGCAATTTCTTTGAGTAAAGTTTGATTTCCAGTAGTACCTGT	98816
QY	8941	CCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTAAATGATGACACAGAGCTGGAG	9000
Db	98815	CCATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTAAATGATGACACAGAGCTGGAG	98756
QY	9001	CATAAATACAATCTCTGACGTTTCATAAAACTGTCGAAAAAGTCTTGTGCTTGAAGTCA	9060
Db	98755	CATAAATACAATCTCTGACGTTTCATAAAACTGTCGAAAAAGTCTTGTGCTTGAAGTCA	98696
QY	9061	CCTACCGGATGCCCGCAGCGTGTGCGGAAAAAGTGTGCCCTCCCTCCCTAGTTTCTGT	9120
Db	98695	CCTACCGGATGCCCGCAGCGTGTGCGGAAAAAGTGTGCCCTCCCTCCCTAGTTTCTGT	98636
QY	9121	AAACCCNAGCCCAAGTCTGCTTTCCCGCGCAGTCCAAATCCCTGCCATCTCGAAGGG	9180
Db	98635	AAACCCNAGCCCAAGTCTGCTTTCCCGCGCAGTCCAAATCCCTGCCATCTCGAAGGG	98576
QY	9181	AAGCAGAGCCCTCTGGGCGCAGCGCAAGGCCCTCTGACTTTCAGGGATAGACTTAGCACT	9240
Db	98575	AAGCAGAGCCCTCTGGGCGCAGCGCAAGGCCCTCTGACTTTCAGGGATAGACTTAGCACT	98516
QY	9241	TTAGCCCCAAGTAACTCTGAAGTCCCAAGACACAGCAAGTGTGGGGTCCAAAGGGCC	9300
Db	98515	TTAGCCCCAAGTAACTCTGAAGTCCCAAGACACAGCAAGTGTGGGGTCCAAAGGGCC	98456
QY	9301	GCACACGAGCAACAGCAATCTGAGATGTTTCTTAAACAGTGTTCCTCCCTGACCGAT	9360
Db	98455	GCACACGAGCAACAGCAATCTGAGATGTTTCTTAAACAGTGTTCCTCCCTGACCGAT	98396
QY	9361	AAGCAAAAAGACCCGAGACAAAATTTGAAACCTCTTCCAGTAGCTCTTCTCAGCCACC	9420
Db	98395	AAGCAAAAAGACCCGAGACAAAATTTGAAACCTCTTCCAGTAGCTCTTCTCAGCCACC	98336
QY	9421	CTCGCAGCAGTAAACATCAATGGTTCCATCGACTACCCCGCAAGAACGACAGCCCGTGG	9480
Db	98335	CTCGCAGCAGTAAACATCAATGGTTCCATCGACTACCCCGCAAGAACGACAGCCCGTGG	98276
QY	9481	GCACCTCCGGGAGAGACTATTTCTGTATCGAGTGCAGCAATCTGCAGCAGATTT	9540
Db	98275	GCACCTCCGGGAGAGACTATTTCTGTATCGAGTGCAGCAATCTGCAGCAGATTT	98216
QY	9541	GGTGAGCCCTTCCAAAAAGACTGAAGTCCAGCGTGTTCCTTTCAGCGTTGACAGCCGCC	9600
Db	98215	GGTGAGCCCTTCCAAAAAGACTGAAGTCCAGCGTGTTCCTTTCAGCGTTGACAGCCGCC	98156
QY	9601	GGGGCCAAATACAGAGAGGCTATGACCTTCCCAAGTACCATATGTCAGAGGCATCA	9660
Db	98155	GGGGCCAAATACAGAGAGGCTATGACCTTCCCAAGTACCATATGTCAGAGGCATCA	98096
QY	9661	TCACCTGTTACCGCAGGACTGTGTATCCGTCCGAGGCGCTCCCTCCCAACCAAGGTTTC	9720
Db	98095	TCACCTGTTACCGCAGGACTGTGTATCCGTCCGAGGCGCTCCCTCCCAACCAAGGTTTC	98036
QY	9721	CTGAGCTCCAGCGAGGTGCATTTCTCCAAATGTGCTGACTGTTTCAGAGCCCTATGTTGCC	9780
Db	98035	CTGAGCTCCAGCGAGGTGCATTTCTCCAAATGTGCTGACTGTTTCAGAGCCCTATGTTGCC	97976
QY	9781	TCCGGGCCATTTACATTTGTGTGCTGTGCTGTAGTCCAGCATCCAGCTCGAGCTTAGAA	9840
Db	97975	TCCGGGCCATTTACATTTGTGTGCTGTGCTGTAGTCCAGCATCCAGCTCGAGCTTAGAA	97916
QY	9841	GGTATTGCAATGAGGGCGCTCGTGTGTTAAATGCTGCTACAGTGATTAAATAGCTAATCA	9900
Db	97915	GGTATTGCAATGAGGGCGCTCGTGTGTTAAATGCTGCTACAGTGATTAAATAGCTAATCA	97856
QY	9901	GGCATTCTCAGTGGAGATGGTACCACCTCCCAAGGGTGGGGGTAGGCAGCAGAGTTCT	9960
Db	97855	GGCATTCTCAGTGGAGATGGTACCACCTCCCAAGGGTGGGGGTAGGCAGCAGAGTTCT	97796
QY	9961	TGGGGTCCACAGAGAGAAAGCATTTCTTATAGTACGCGAGTGG 10000	
Db	97795	TGGGGTCCACAGAGAGAAAGCATTTCTTATAGTACGCGAGTGG 97756	
RESULT 3			
AL157838/c			
LOCUS			
DEFINITION			
Human DNA sequence from clone RP4-724E16 on chromosome			
20q13.12-13.32 Contains the ZNF217 gene for zinc finger protein			
217, a novel gene, a putative novel gene, ESTs, GSSs, STSs and two			
CpG islands, complete sequence.			
ACCESSION			
AL157838			
VERSION			
AL157838.24 GI:9588158			
KEYWORDS			
HTG; CpG island; zinc finger; ZNF217.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1 (bases 1 to 128871)			
AUTHORS			
Wilson, S.			
TITLE			
Direct Submission			
JOURNAL			
Submitted (23-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,			
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk			
COMMENTS			
requests: clonerequest@sanger.ac.uk			
On Jul 31, 2000 this sequence version replaced gi:9408255.			
During sequence assembly data is compared from overlapping clones.			
Where differences are found these are annotated as variations			
together with a note of the overlapping clone name. Note that the			
variation annotation may not be found in the sequence submission			
corresponding to the overlapping clone, as we submit sequences with			
only a small overlap as described above.			
The following abbreviations are used to associate primary accession			
numbers given in the feature table with their source databases:			
Em: ENBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information			
on the WORMPEP database can be found at			
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence			
was generated from part of bacterial clone contigs of human			
Chromosome 20, constructed by the Sanger Centre Chromosome 20			
Mapping Group. Further information can be found at			
http://www.sanger.ac.uk/HGP/Chr20			
This sequence is the entire insert of clone RP4-724E16 The true			
right end of clone RP5-823G15 is at 19684 in this sequence. This			
sequence was finished as follows unless otherwise noted: all			
regions were either double-stranded or sequenced with an alternate			
chemistry or covered by high quality data (i.e., phred quality >=			
30); an attempt was made to resolve all sequencing problems, such			
as compressions and repeats; all regions were covered by at least			
one plasmid subclone or more than one M13 subclone, and the			
assembly was confirmed by restriction digest. RP4-724E16 is from			
the library RPCI-4 constructed by the group of Pieter de Jong. For			
further details see			
http://www.ckori.org/bacpac/home.htm			
VECTOR: PCYPAC2.			
FEATURES			
Location/Qualifiers			
1..128871			
/organism="Homo sapiens"			
/mol_type="genomic DNA"			
/db_xref="taxon:9606"			
/chromosome="20"			
/map="q13.12-13.32"			
/clone="RP4-724E16"			
/clone_lib="RPCI-4"			
repeat_region 1..54			
/note="MLT2A repeat: matches 362. .415 of consensus"			
repeat_region 55..368			

repeat_region /note="AluX repeat: matches 1. .312 of consensus"
369. .559
/note="AluSq repeat: matches 126. .312 of consensus"
1137. .1270
/note="FLAM C repeat: matches 1. .133 of consensus"
complement(1396. .1840)
misc_feature /note="match: GSS: Em:AQ261624"
1415. .1709
repeat_region /note="AluJo repeat: matches 14. .288 of consensus"
1744. .1902
repeat_region /note="MIR repeat: matches 14. .175 of consensus"
1871. .2326
misc_feature /note="match: GSS: Em:AQ673688"
1948. .2041
repeat_region /note="MER5A repeat: matches 6. .101 of consensus"
2622. .2787
repeat_region /note="MER5A repeat: matches 4. .189 of consensus"
2806. .2870
repeat_region /note="MER5A repeat: matches 10. .75 of consensus"
2872. .3026
repeat_region /note="L1MA4 repeat: matches 6017. .6162 of consensus"
3027. .3307
repeat_region /note="AluJo repeat: matches 1. .290 of consensus"
3308. .3436
repeat_region /note="L1MA4 repeat: matches 6162. .6293 of consensus"
3455. .3551
repeat_region /note="MER5A repeat: matches 61. .183 of consensus"
3564. .3697
repeat_region /note="MER5B repeat: matches 14. .146 of consensus"
3871. .4004
repeat_region /note="AluSq/x repeat: matches 1. .134 of consensus"
4249. .4274
repeat_region /note="13 copies 2 mer gt 96% conserved"
4810. .5101
repeat_region /note="AluX repeat: matches 1. .297 of consensus"
5110. .5413
repeat_region /note="AluSc repeat: matches 1. .305 of consensus"
5492. .5648
repeat_region /note="L2 repeat: matches 2589. .2745 of consensus"
5659. .6135
repeat_region /note="MER1A repeat: matches 1. .488 of consensus"
6164. .6225
repeat_region /note="31 copies 2 mer aa 71% conserved"
6245. .6379
repeat_region /note="AluSq/x repeat: matches 162. .296 of consensus"
6380. .6537
repeat_region /note="79 copies 2 mer aa 60% conserved"
6558. .6967
repeat_region /note="AluSq repeat: matches 1. .309 of consensus"
7320. .7542
repeat_region /note="MER50B repeat: matches 117. .341 of consensus"
7545. .7838
repeat_region /note="AluX repeat: matches 1. .293 of consensus"
8465. .9068
misc_feature /note="match: GSS: Em:AQ386102"
8523. .8814
repeat_region /note="AluJb repeat: matches 1. .294 of consensus"
9262. .9561
repeat_region /note="AluSq repeat: matches 1. .301 of consensus"
9562. .9599
repeat_region /note="19 copies 2 mer tt 81% conserved"
9606. .9733
repeat_region /note="MER33 repeat: matches 17. .148 of consensus"
9761. .9915
repeat_region /note="MER5A repeat: matches 4. .183 of consensus"
9980. .10108
repeat_region /note="L2 repeat: matches 2622. .2747 of consensus"
10128. .10437
repeat_region /note="AluX repeat: matches 2. .312 of consensus"
10654. .10724
repeat_region /note="MIR repeat: matches 63. .137 of consensus"
10781. .11083
repeat_region /note="AluX repeat: matches 1. .306 of consensus"

misc_feature complement(10897. .11319)
/note="match: GSS: Em:AQ495673"
misc_feature complement(10991. .11319)
/note="match: GSS: Em:AQ283579"
11462. .11759
repeat_region /note="AluSq repeat: matches 1. .299 of consensus"
11878. .12182
repeat_region /note="AluX repeat: matches 1. .300 of consensus"
12183. .12587
repeat_region /note="L2 repeat: matches 5757. .6163 of consensus"
13333. .13637
repeat_region /note="MLTIG repeat: matches 3. .338 of consensus"
14201. .14507
repeat_region /note="AluY repeat: matches 1. .308 of consensus"
14540. .14694
repeat_region /note="L2 repeat: matches 2550. .2709 of consensus"
14764. .14868
repeat_region /note="MIR repeat: matches 103. .208 of consensus"
14891. .15152
repeat_region /note="AluJo repeat: matches 12. .299 of consensus"
15162. .15293
repeat_region /note="L2 repeat: matches 1697. .1824 of consensus"
15317. .15544
repeat_region /note="L1MB3 repeat: matches 5943. .6173 of consensus"
16306. .16603
repeat_region /note="AluSc repeat: matches 1. .306 of consensus"
16610. .16901
repeat_region /note="AluX repeat: matches 1. .292 of consensus"
16989. .17035
repeat_region /note="MLTIC repeat: matches 420. .463 of consensus"
17036. .17338
repeat_region /note="AluJo repeat: matches 1. .301 of consensus"
17339. .17730
repeat_region /note="MLTIC repeat: matches 1. .420 of consensus"
17958. .18269
repeat_region /note="AluX repeat: matches 1. .311 of consensus"
18288. .18315
repeat_region /note="MER91 repeat: matches 32. .59 of consensus"
18383. .18507
repeat_region /note="MIR repeat: matches 23. .148 of consensus"
18544. .18650
repeat_region /note="MIR repeat: matches 137. .261 of consensus"
18799. .18820
repeat_region /note="11 copies 2 mer aa 100% conserved"
18835. .19184
repeat_region /note="THE1A repeat: matches 1. .354 of consensus"
19231. .19522
misc_feature /note="match: STS: Em:HSAL53VF5"
19297. .19346
repeat_region /note="25 copies 2 mer ca 100% conserved"
19426. .19584
repeat_region /note="MIR repeat: matches 6. .147 of consensus"
19624. .19916
repeat_region /note="AluX repeat: matches 1. .295 of consensus"
20005. .20204
repeat_region /note="MER20 repeat: matches 1. .218 of consensus"
20215. .20317
repeat_region /note="MER5A repeat: matches 4. .108 of consensus"
20410. .20541
repeat_region /note="MER5B repeat: matches 1. .120 of consensus"
20544. .20687
repeat_region /note="AluSq/x repeat: matches 154. .297 of consensus"
20689. .20814
repeat_region /note="FLAM A repeat: matches 14. .136 of consensus"
20816. .21218
repeat_region /note="Charlie4a repeat: matches 17. .427 of consensus"
21337. .21634
repeat_region /note="AluY repeat: matches 1. .297 of consensus"

Query Match 99.8%; Score 9980.6; DB 9; Length 128871;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 9996; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

QY 1 CCATCATATTTCTTATTTTTTGGCGGAGAGGGAGACTTCTGTTGCCAGGCTGG 60
Db |||||
74006 CCATCATATTTCTTATTTTTTGGCGGAGAGGGAGACTTCTGTTGCCAGGCTGG 73947
QY 61 ACCAGTGGTGGCATCTTGGCTCACTGCAACCTCCACCTCTGGGTCAAGTGATTTCCCAA 120
Db |||||
73946 ACCAGTGGTGGCATCTTGGCTCACTGCAACCTCCACCTCTGGGTCAAGTGATTTCCCAA 73887
QY 121 ATAGCTGGGATTTACAGGTGTGTATTAACATGCCAGCTAAATTTTTGTATTTTAGAGAT 180
Db |||||
73886 ATAGCTGGGATTTACAGGTGTGTATTAACATGCCAGCTAAATTTTTGTATTTTAGAGAT 73827
QY 181 RAGGGSTTTTCCACATGTTGGCCAGGCTGGTCTCCAACTCCTGGCTCATGTGATCCACC 240
Db |||||
73826 RAGGGSTTTTCCACATGTTGGCCAGGCTGGTCTCCAACTCCTGGCTCATGTGATCCACC 73767
QY 241 ACTTCGGCTTCCCAAAGCATTTGGGAGTATAGGTGTGAGCCACTATACCCGCTCCTCACATC 300
Db |||||
73766 ACTTCGGCTTCCCAAAGCATTTGGGAGTATAGGTGTGAGCCACTATACCCGCTCCTCACATC 73707
QY 301 ATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTCTAAGGATGTCAGTAGAG 360
Db |||||
73706 ATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTCTAAGGATGTCAGTAGAG 73647
QY 361 RAGTGGAGTTCCCAAATTTACAGTTTACAGTTTACAGTTTACAGTTTACAGTTTACAGTTT 420
Db |||||
73646 RAGTGGAGTTTCCCAAATTTACAGTTTACAGTTTACAGTTTACAGTTTACAGTTTACAGTTT 73587
QY 421 AATGTTGAGAGCTGACATAGGAGCTAACTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
Db |||||
73586 AATGTTGAGAGCTGACATAGGAGCTAACTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 73527
QY 481 TCTCACTGAACCTTGAATTTGCTAAATAAGGACATTTAAAGGACATTTAAAGGACATTTAAAGG 540
Db |||||
73526 TCTCACTGAACCTTGAATTTGCTAAATAAGGACATTTAAAGGACATTTAAAGGACATTTAAAGG 73467
QY 541 TATTCGCTATTGCCACTATTTGATTTTAAAGGACATTTAAAGGACATTTAAAGGACATTTAAAGG 600
Db |||||
73466 TATTCGCTATTGCCACTATTTGATTTTAAAGGACATTTAAAGGACATTTAAAGGACATTTAAAGG 73407
QY 601 AGGAAGGACCTCAATTAATGAGTCTTTGTTCTGGCCAGGGAAACAGCGTTGTGAGAA 660
Db |||||
73406 AGGAAGGACCTCAATTAATGAGTCTTTGTTCTGGCCAGGGAAACAGCGTTGTGAGAA 73347
QY 661 TTTGATTAACGTTTTTCTAGGGTATGTGCTGTTATTTAGTTTAAACCTTGGCTGGGACGC 720
Db |||||
73346 TTTGATTAACGTTTTTCTAGGGTATGTGCTGTTATTTAGTTTAAACCTTGGCTGGGACGC 73287
QY 721 TAGCATTCAAGTAAATACTTTGTAATAGCAATGAACTTAAGCTTCTATGATAGAA 780
Db |||||
73286 TAGCATTCAAGTAAATACTTTGTAATAGCAATGAACTTAAAGCTTCTATGATAGAA 73227
QY 781 CCTAAGTCACTTCAATTTCTGATTAGCAGAGTAAATGAAATTTCTTTTCAATGTGTAGCT 840
Db |||||
73226 CCTAAGTCACTTCAATTTCTGATTAGCAGAGTAAATGAAATTTCTTTCAATGTGTAGCT 73167
QY 841 CTATCCCGAGAACCAAGATATTTGAACTGTAAAGGCCATCTATAGTTTAAACCAACTG 900
Db |||||
73166 CTATCCCGAGAACCAAGATATTTGAACTGTAAAGGCCATCTATAGTTTAAACCAACTG 73107
QY 901 CGTTAAATAGATAATAGAAAGATGTGTATGTGGCAGTGCACAACTTGAAGGTTGTGACTA 960
Db |||||
73106 CGTTAAATAGATAATAGAAAGATGTGTATGTGGCAGTGCACAACTTGAAGGTTGTGACTA 73047
QY 961 GAACTCGGGTCTCTGGAGTGTCTTATTTATCACCAAGCTGGTCACCAGGCCATGTGT 1020
Db |||||
73046 GAACTCGGGTCTCTGGAGTGTCTTATTTATCACCAAGCTGGTCACCAGGCCATGTGT 72987
QY 1021 TGATCCTCAATGTGTATAGCAACAAAGAAAGACTTTCAGGACATTTCTTTCTTTACCCCTA 1080
Db |||||
72986 TGATCCTCAATGTGTATAGCAACAAAGAAAGACTTTCAGGACATTTCTTTCTTTACCCCTA 72927
QY 1081 ATCCTTGATCTGAGTCTTTATTTAGAAAGCTTAAATGTTTAAAGATCTAGTTTATTCAAA 1140

72926 ATCCTTGATCTGAGTCTTTATTTAGAAAGCTTAAATGTTTAAAGATCTAGTTTATTCAAA 72867
QY |||||
1141 CTAAAGATAACAAGGAGTATGAGAAATTTCTATTTCCGAGTGTAAAGGAGGAGATGTTTCC 1200
Db |||||
72866 CTAAAGATAACAAGGAGTATGAGAAATTTCTATTTCCGAGTGTAAAGGAGGAGATGTTTCC 72807
QY 1201 TTGGCTTCTCTGAGCTCAGGCCCTTCTGCTCTTTAAGNAAGTAGAGAGGAGGAA 1260
Db |||||
72806 TTGGCTTCTCTGAGCTCAGGCCCTTCTGCTCTTTAAGNAAGTAGAGAGGAGGAA 72747
QY 1261 AGTAAAGTATGCTTTTGTTTTTTAAAGTTTACTTTCTCGGAGTAGTTTGCATGCTTTTG 1320
Db |||||
72746 AGTAAAGTATGCTTTTGTTTTTTAAAGTTTACTTTCTCGGAGTAGTTTGCATGCTTTTG 72687
QY 1321 GTTTCTTTGGGTGGAAATTAAGTCTTAAGTTTTTAAAGTAGTTTGGGACTATTTTAAACAA 1380
Db |||||
72686 GTTTCTTTGGGTGGAAATTAAGTCTTAAGTTTTTAAAGTAGTTTGGGACTATTTTAAACAA 72627
QY 1381 TGCCTATCCAATGTTTGCATAAAGGCAGAGGTTATTTGGCTTTAGAGTTAATTTCTCTC 1440
Db |||||
72626 TGCCTATCCAATGTTTGCATAAAGGCAGAGGTTATTTGGCTTTAGAGTTAATTTCTCTC 72567
QY 1441 CAGGAGTGAAATTAAGTCTTTAAACCAAGCAGCAGAGCTAAATAAAGTAAATTTTCCAC 1500
Db |||||
72566 CAGGAGTGAAATTAAGTCTTTAAACCAAGCAGCAGAGCTAAATAAAGTAAATTTTCCAC 72507
QY 1501 CTGGCCAGTGATGATGAAAGGTAGATTTAAAAAATAGAGGCCCAATTTTCTGATGA 1560
Db |||||
72506 CTGGCCAGTGATGATGAAAGGTAGATTTAAAAAATAGAGGCCCAATTTTCTGATGA 72447
QY 1561 RAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTAAATCTATACATTCAC 1620
Db |||||
72446 RAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTAAATCTATACATTCAC 72387
QY 1621 AAAGGAGCTTTGTTGATGTTCTTTCCCTATTTGTTGTTGGACTAGGAAGCCCAACCACT 1680
Db |||||
72386 AAAGGAGCTTTGTTGATGTTCTTTCCCTATTTGTTGTTGGACTAGGAAGCCCAACCACT 72327
QY 1681 GCTTTTGAAGGCAGAAAGTCTGTGAAAGCAGCTGGGATTTGAAACAGTGGATTTGAGGTT 1740
Db |||||
72326 GCTTTTGAAGGCAGAAAGTCTGTGAAAGCAGCTGGGATTTGAAACAGTGGATTTGAGGTT 72267
QY 1741 TCGAATATCCAGTGAAACCAATATATCAGGGTCCCTGGCCCAAGATGAGGACCATTC 1800
Db |||||
72266 TCGAATATCCAGTGAAACCAATATATCAGGGTCCCTGGCCCAAGATGAGGACCATTC 72207
QY 1801 TGAGGTGTTAAGTATTTCTTGAATGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 1860
Db |||||
72206 TGAGGTGTTAAGTATTTCTTGAATGGGATTTTAGGAAAAGTTTCTGTATTTCTGTGCTC 72147
QY 1861 ATTTTGTGACCTCTGTATGTGCAAAAATCTCTAAGGGGTGTTGGGCACTTAGATTTCT 1920
Db |||||
72146 ATTTTGTGACCTCTGTATGTGCAAAAATCTCTAAGGGGTGTTGGGCACTTAGATTTCT 72087
QY 1921 TGGATGCAAGATTTGTTGATATGAAAACAAATTTTAAATGTTTGTATACACTGGATTT 1980
Db |||||
72086 TGGATGCAAGATTTGTTGATATGAAAACAAATTTTAAATGTTTGTATACACTGGATTT 72027
QY 1981 AAAATAGTTTACTAAAGTGTTTTAAATTTTCAATTTTCAATTTTCAAGTTCTTATAGTCT 2040
Db |||||
72026 AAAATAGTTTACTAAAGTGTTTTAAATTTTCAATTTTCAATTTTCAAGTTCTTATAGTCT 71967
QY 2041 TTAGATTTAGGAGGCTGTGTAGTGCAATCCCATGTCGATTTTGTGGCACTTAAAGT 2100
Db |||||
71966 TTAGATTTAGGAGGCTGTGTAGTGCAATCCCATGTCGATTTTGTGGCACTTAAAGT 71907
QY 2101 ATTCACTCAATTTTAAACAAATTTCTGACCTTAAACCTTGAATTTTAAAGTTCGATAA 2160
Db |||||
71906 ATTCACTCAATTTTAAACAAATTTCTGACCTTAAACCTTGAATTTTAAAGTTCGATAA 71847
QY 2161 AGCACTGATTTAAACTGGATTTTAACTGGATGAAATTTCTGATTTTAAAGTGTACTGACT 2220
Db |||||

QY 4381 TCACAAAACACACTTTTAAAAAATAAATTAACCTTGTGCATCCAGCCCAATGCACTGCTCT 4440
DB |||||
DB 69626 TCACAAAACACACTTTTAAAAAATAAATTAACCTTGTGCATCCAGCCCAATGCACTGCTCT 69567
QY 4441 TAACTGGGCGGATTTTGTTCCTCAATCAGTATCTGGCAATGCTCTGAGGCAATTTTGGTGT 4500
DB |||||
DB 69566 TAACTGGGCGGATTTTGTTCCTCAATCAGTATCTGGCAATGCTCTGAGGCAATTTTGGTGT 69507
QY 4501 CATACTGTGTGTGGGTGTGCTGTGCTGTGCTATCCAGTGGGCGAGAGCCAGGACACTGCT 4560
DB |||||
DB 69506 CATACTGTGTGTGGGTGTGCTGTGCTATCCAGTGGGCGAGAGCCAGGACACTGCT 69447
QY 4561 CAGCATGTACAGTGCACAGGACAGCCCATCATCAAGAATTAATCTGGTCCCAATGTC 4620
DB |||||
DB 69446 CAGCATGTACAGTGCACAGGACAGCCCATCATCAAGAATTAATCTGGTCCCAATGTC 69387
QY 4621 AATAGTTTGGAGCAATTTGAGAGACCTTAGCCCTTCACTTAAAGTTTCTGGCGTTTCTGTATCT 4680
DB |||||
DB 69386 AATAGTTTGGAGCAATTTGAGAGACCTTAGCCCTTCACTTAAAGTTTCTGGCGTTTCTGTATCT 69327
QY 4681 TTTTCTGTAGTGAATTTTCTAGTGGCCATAAAAGGTACTGGGAG---TGATCAACTAGAGC 4737
DB |||||
DB 69326 TTTTCTGTAGTGAATTTTCTAGTGGCCATAAAAGGTACTGGGAGGTATGATCAACTAGAGC 69267
QY 4738 CAGGAATATATTTGGGCGAGCGTTTGGTGTCTGCTCCAAACCTTCTCTTCTGTCTGGC 4797
DB |||||
DB 69266 CAGGAATATATTTGGGCGAGCGTTTGGTGTCTGCTCCAAACCTTCTCTTCTGTCTGGC 69207
QY 4798 AAGCTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTGCATAAAATACAAGG 4857
DB |||||
DB 69206 AAGCTAGTATCCATTTATAGGTACCTCAGGAACCCAAATGATTTGTGCATAAAATACAAGG 69147
QY 4858 AATGTGAGCACACTGAAGACATTTTAAAGAGGCTCATTTGCTCAGCAGAAATTTTCAGTG 4917
DB |||||
DB 69146 AATGTGAGCACACTGAAGACATTTTAAAGAGGCTCATTTGCTCAGCAGAAATTTTCAGTG 69087
QY 4918 TACTAGTGGCATTTATAGAAAGAGAGGTGATCAGTGAAGGCACTGCTCACATAATATTC 4977
DB |||||
DB 69086 TACTAGTGGCATTTATAGAAAGAGAGGTGATCAGTGAAGGCACTGCTCACATAATATTC 69027
QY 4978 TGAGCCCTGTGGGGGTTTATCTAGGCGAAAGGATTCACCTGTGTTGGAGTTGGGCCCA 5037
DB |||||
DB 69026 TGAGCCCTGTGGGGGTTTATCTAGGCGAAAGGATTCACCTGTGTTGGAGTTGGCCCA 68967
QY 5038 TCCTCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTGTATTTGTTTGAATAGAGATC 5097
DB |||||
DB 68966 TCCTCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTGTATTTTGTGAAATAGAGATC 68907
QY 5098 TTGCTGCTTAAACAGTTGAAAGACCTGTATGGGCGAGCCGTGAATTGACAGCGAATGA 5157
DB |||||
DB 68906 TTGCTGCTTAAACAGTTGAAAGACCTGTATGGGCGAGCCGTGAATTGACAGCGAATGA 68847
QY 5158 TGGGAACATGAATCCGCTCTTAGGGAAGCATCTGTCAAAGTGGTCTTGGTTAAAAACAAGT 5217
DB |||||
DB 68846 TGGGAACATGAATCCGCTCTTAGGGAAGCATCTGTCAAAGTGGTCTTGGTTAAAAACAAGT 68787
QY 5218 GCCTCCTCTCCTCAGGTGCATTTGATGTGTGCTGAAATTCMTGGAAAACTGGGTGTAT 5277
DB |||||
DB 68786 GCCTCCTCTCCTCAGGTGCATTTGATGTGTGCTGAAATTCMTGGAAAACTGGGTGTAT 68727
QY 5278 GAGACCCAGATGAATTTGCCACACAGGATTTGATTTGAGCTTCTTCCCTCACCTGCTCTCAG 5337
DB |||||
DB 68726 GAGACCCAGATGAATTTGCCACACAGGATTTGATTTGAGCTTCTTCCCTCACCTGCTCTCAG 68667
QY 5338 CCAGTGGCCAGTTTCTTCTGATCATGTGATTTGAAGTGAAGAACTGTAGTCTGTATATCAA 5397
DB |||||
DB 68666 CCAGTGGCCAGTTTCTTCTGATCATGTGATTTGAAGTGAAGAACTGTAGTCTGTATCAA 68607
QY 5398 ATCTTTAGAAATGTTTGTAGTTTCTGGGACACAGGAAACCCAGCACTTTAGCATACTACA 5457
DB |||||
DB 68606 ATCTTTAGAAATGTTTGTAGTTTCTGGGACACAGGAAACCCAGCACTTTAGCATACTACA 68547
QY 5458 AATCTAAATGCTTAATGGGCATATAAAAGAGGCTTTTAAACACAGACTCTCAGTTAGCTAA 5517

DB |||||
QY 68546 AATCTAATGCTTAATGCAATCATAAAAGAGGCTTTTAAACACAGACTCCAGTTAGCTAA 68487
QY 5518 GTGGTTTCTGCTAGTGC CGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTGA 5577
DB |||||
DB 68486 GTGGTTTCTGCTAGTGC CGGTACTGTTGCAGGGGCCCTGTGAGATGCCCCAGTTCCCTGA 68427
QY 5578 AAGAAATGAAAAGGCGCAGTTACCGGTAGTGTGGTGTGGAAAACATGGGCTAGATCATCAGG 5637
DB |||||
DB 68426 AAGAAATGAAAAGGCGCAGTTACCGGTAGTGTGGTGTGGAAAACATGGGCTAGATCATCAGG 68367
QY 5638 CAGGACGAATGCTGCTGCTGGGTGGGAGCACCCAGCTTGGCGCTGTGAGTTCTGCTTCT 5697
DB |||||
DB 68366 CAGGACGAATGCTGCTGCTGGGTGGGAGCACCCAGAGTTGGCGCTGTGAGTTCTGCTTCT 68307
QY 5698 ACCACTGGGTTGTTTTGTGACCAATTAATGAGTTGCTTAACTTTCTTTGCTACTATTTCC 5757
DB |||||
DB 68306 ACCACTGGGTTGTTTTGTGACCAATTAATGAGTTGCTTAACTTTCTTTGCTACTATTTCC 68247
QY 5758 CTGTTTGCAAAATGCTTCAITTGACCCCTGCTTCCACCTCCCAAGGACAAATTTCAACAGC 5817
DB |||||
DB 68246 CTGTTTGCAAAATGCTTCAITTGACCCCTGCTTCCACCTCCCAAGGACAAATTTCAACAGC 68187
QY 5818 CTATTTGTAAAAAGATCACAGTCTTTAAAAAATAAATGCTGTAAGTCAAGAGTGTATGCT 5877
DB |||||
DB 68186 CTATTTGTAAAAAGATCACAGTCTTTAAAAAATAAATGCTGTAAGTCAAGAGTGTATGCT 68127
QY 5878 TGAAGAGCAGGAACACAGGTAGATGTGAAATGTGATGCTCTTTGTTCTTAAAGAAAGGC 5937
DB |||||
DB 68126 TGAAGAGCAGGAACACAGGTAGATGTGAAATGTGATGCTCTTTGTTCTTAAAGAAAGGC 68067
QY 5938 ATTTTCATAGCTTTTGGATATGACCAACATACATAATCCTGACACATAGTTGGGAGT 5997
DB |||||
DB 68066 ATTTTCATAGCTTTTGGATATGACCAACATACATAATCCTGACATAGTTGGGAGT 68007
QY 5998 CGSAAATTTGCAACAGCCAGTTATAAACCCAGCTAGTTTGGGTATGATTTCTTAAGAAAA 6057
DB |||||
DB 68006 CGSAAATTTGCAACAGCCAGTTATAAACCCAGCTAGTTTGGGTATGATTTCTTAAGAAAA 67947
QY 6058 AAAAGCTGGCCATCTGTATTTTGGGAAATGATTTTCCCTTAAACTTATATTTCTTAGTAG 6117
DB |||||
DB 67946 AAAAGCTGGCCATCTGTATTTTGGGAAATGATTTTCCCTTAAACTTATATTTCTTAGTAG 67887
QY 6118 TCTAGATTTATCATATTTGTAATCTATCATCTGCTGCTTTTAAAGCTTTAAGAACTCAAGTA 6177
DB |||||
DB 67886 TCTAGATTTATCATATTTGTAATCTATCATCTGCTGCTTTTAAAGCTTTAAGAACTCAAGTA 67827
QY 6178 AATTTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGCTGTCTCTTTACAGGTG 6237
DB |||||
DB 67826 AATTTTTTTTTCTTTCTTTAGACACTATATAGATCATCAAGGCTGTCTCTTTACAGGTG 67767
QY 6238 GATAGTATATGATCTACAGTGAGGGGACATTTTAAAACTTAAAACTTCAATGTTGTTT 6297
DB |||||
DB 67766 GATAGTATATGATCTACAGTGAGGGGACATTTTAAAACTTAAAACTTCAATGTTGTTT 67707
QY 6298 TGGGGGTGGTATTTTAAACGGCAGACCTCTGATTTGCTTTTGGAGGCTGTGTGTGTTT 6357
DB |||||
DB 67706 TGGGGGTGGTATTTTAAACGGCAGACCTCTGATTTGCTTTTGGAGGCTGTGTGTGTTT 67647
QY 6358 GAAGTTCCTGCTCCTTCCAGTGGGACTTAACTTCTCTGATGCGAGTGCAGACACATTTGT 6417
DB |||||
DB 67646 GAAGTTCCTGCTCCTTCCAGTGGGACTTAACTTCTCTGATGCGAGTGCAGACACATTTGT 67587
QY 6418 CCTATGTCCTGCAGAAACTAAAGCCAAACACTGTGCATCTGGGACAGGTTTTTCAATTTGT 6477
DB |||||
DB 67586 CCTATGTCCTGCAGAAACTAAAGCCAAACACTGTGCATTTGGGGACAGGTTTTTCAATTTGT 67527
QY 6478 CAGATCTCTTTGCGCCACATGAGTCTTTGTGACAAATACAGCTCTGTTTTCCAAAACTTTG 6537
DB |||||
DB 67526 CAGATCTCTTTGCGCCACATGAGTGTTTGTGACAAATACAGCTCTGTTTTCCAAAACTTTG 67467
QY 6538 CTAAATTTTGACAGACTTTTCTTAGGTGCTTGGCCAAATGCCAGACTTTTCTTTCTGTGAA 6597

QY 8758 GATGCGAGTACCACCAATACCTTGAAGTTAGTCCCAAGAGAAACGAGAGCCGA 8817
Db 65246 GATGCGAGTACCACCAATACCTTGAAGTTAGTCCCAAGAGAAACGAGAGCCGA 65187
QY 8818 GCTGACTGCAGATACAGGCAAGTGTGAGTTGTACGAAAAACCTTTTAAATTTATCCGTG 8877
Db 65186 GCTGACTGCAGATACAGGCAAGTGTGAGTTGTACGAAAAACCTTTTAAATTTATCCGTG 65127
QY 8878 GGGGCTCTTACAATTTGCGCGCAATTTCTTTGAGTAAAGTTTGATTTCCAAAGTATCACC 8937
Db 65126 GGGGCTCTTACAATTTGCGCGCAATTTCTTTGAGTAAAGTTTGATTTCCAAAGTATCACC 65067
QY 8938 TGTTCATTTTGTACCTTCAAGACATTTTATCCAGAACTTTTAAATATGATGCACAGAGCTG 8997
Db 65066 TGTTCATTTTGTACCTTCAAGACATTTTATCCAGAACTTTTAAATATGATGCACAGAGCTG 65007
QY 8998 GAGCATAAATACAATCTCTGACGTTTCATAAAACTGTGCAAAACAAAGTCTTGTCTAGAAAGT 9057
Db 65006 GAGCATAAATACAATCTCTGACGTTTCATAAAACTGTGCAAAACAAAGTCTTGTCTAGAAAGT 64947
QY 9058 CGACGTACCGGATGCGCGCAAGTGTGCTGGGAAAGATGTGCCCTCCCTCTAGTTTC 9117
Db 64946 CGACGTACCGGATGCGCGCAAGTGTGCTGGGAAAGATGTGCCCTCCCTCTAGTTTC 64887
QY 9118 TGTAACCCAGCCCAAGTCTGCTTCCGCGCAGTCCAAATCCTGCCATCTCGGAAG 9177
Db 64886 TGTAACCCAGCCCAAGTCTGCTTCCGCGCAGTCCAAATCCTGCCATCTCGGAAG 64827
QY 9178 GGGAAAGCAGAGCCCTCTCTGGGCGAGGAGGCGCTCTGACCTTCCAGGATAGACTCTAGC 9237
Db 64826 GGGAAAGCAGAGCCCTCTCTGGGCGAGGAGGCGCTCTGACCTTCCAGGATAGACTCTAGC 64767
QY 9238 ACTTTAGCCCCAAGTAACTGAAGTCCCAAGACACAGCAAGTGTGGGGTCCAAAGG 9297
Db 64766 ACTTTAGCCCCAAGTAACTGAAGTCCCAAGACACAGCAAGTGTGGGGTCCAAAGG 64707
QY 9298 GCGGCCACGAGCAACAGCAATCTGAGATGTTTCTTAAACAGTGTTCCTCCCTGACCG 9357
Db 64706 GCGGCCACGAGCAACAGCAATCTGAGATGTTTCTTAAACAGTGTTCCTCCCTGACCG 64647
QY 9358 GATAAGACAAAAGACCCGAGACAAAATTTGAAACCTTCTTCCAGTAGCTCTTCTCAGGCC 9417
Db 64646 GATAAGACAAAAGACCCGAGACAAAATTTGAAACCTTCTTCCAGTAGCTCTTCTCAGGCC 64587
QY 9418 ACCCTCGGAGAGTAAATCAATAGTTCATCGACTACCCCGCAAGAACGACAGCCG 9477
Db 64586 ACCCTCGGAGAGTAAATCAATAGTTCATCGACTACCCCGCAAGAACGACAGCCG 64527
QY 9478 TGGGCACCTCCGGGAGAGACTATTCTGTAATCGAGTGCAGCAATATCGCAGCAAA 9537
Db 64526 TGGGCACCTCCGGGAGAGACTATTCTGTAATCGAGTGCAGCAATATCGCAGCAAA 64467
QY 9538 TTTGTGTAGCCCCCTTCCAAAAGACTGAAGTCCAGCGTGTGTCCTTTGACCTTGACCCAG 9597
Db 64466 TTTGTGTAGCCCCCTTCCAAAAGACTGAAGTCCAGCGTGTGTCCTTTGACCTTGACCCAG 64407
QY 9598 CCGCGGGCAATTAAGAGAGGCTATGACCTTCCCAAGTACCAATATGCTCAGAGCATC 9657
Db 64406 CCGCGGGCAATTAAGAGAGGCTATGACCTTCCCAAGTACCAATATGCTCAGAGCATC 64347
QY 9658 ACATCACTGTTCACGAGAGTGTGTATCCGTCGAGGCGCTGCTCCCAAAACCAAGG 9717
Db 64346 ACATCACTGTTCACGAGAGTGTGTATCCGTCGAGGCGCTGCTCCCAAAACCAAGG 64287
QY 9718 TTTCTGTAGCTCCAGGAGGCTCGATTTCTCCAAATGTGCTGACTGTTCAGAGCCCTATGCT 9777
Db 64286 TTTCTGTAGCTCCAGGAGGCTCGATTTCTCCAAATGTGCTGACTGTTCAGAGCCCTATGCT 64227
QY 9778 GGCCTCGGGCCACTTTTACCTGTGTGCTGCTGTGAGTCCAGCATTCAGACTCGAGTTA 9837
Db 64226 GGCCTCGGGCCACTTTTACCTGTGTGCTGCTGTGAGTCCAGCATTCAGACTCGAGTTA 64167
QY 9838 GAAGGTATTGTCATGAGGGCGTCTGTTTAAATGGCTGCTACAGTGAATTAATAGCTAAT 9897

Db 64166 GAAGGTATTGTCATGAGGGCGTCTGTTTAAATGGCTGCTACAGTGAATTAATAGCTAAT 64107
QY 9898 CCAGCATTTCTCAGTGGAGATGGTACCACTCCCAAGAGGTGGGGGTAGCCAGCCAGAGT 9957
Db 64106 CCAGCATTTCTCAGTGGAGATGGTACCACTCCCAAGAGGTGGGGGTAGCCAGCCAGAGT 64047
QY 9958 TCTTGGGGTCCACAGAGAGAGCATTTCTTAGATACGGCAGTGG 10000
Db 64046 TCTTGGGGTCCACAGAGAGAGCATTTCTTAGATACGGCAGTGG 64004

RESULT 4
AC116668
LOCUS
DEFINITION Trypanosoma brucei chromosome V clone RPC193-45E22, *** SEQUENCING
IN PROGRESS ***, 3 unordered pieces.
AC116668
ACCESSION
VERSION AC116668.5 GI:20376999
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Trypanosoma brucei
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.
REFERENCE
1 (bases 1 to 105023)
AUTHORS El-Sayed, N.M., Ghedin, E., Song, J., Larkin, C., Wanless, D., Jones, K., Peterson, J., Hou, L., Zhao, H., Zhou, H., Miltscher, J., Pai, G., Van Aken, S., Utterback, T., Khalak, H.G., Gerard, C., Leech, V., Ullu, E., Melville, S., White, O., Adams, M.D., Donelson, J.B. and Fraser, C.M.
Trypanosoma brucei GUTat10.1 RPC193-45E22 BAC genomic sequence
Unpublished
2 (bases 1 to 105023)
AUTHORS El-Sayed, N.M., Khalak, H. and Adams, M.D.
Direct Submission
TITLE Submitted (02-APR-2002) The Institute for Genomic Research, 9712
JOURNAL Medical Center Dr, Rockville, MD 20850, USA
3 (bases 1 to 105023)
AUTHORS El-Sayed, N.M., Khalak, H. and Adams, M.D.
Direct Submission
TITLE Submitted (01-MAY-2002) The Institute for Genomic Research, 9712
JOURNAL Medical Center Dr, Rockville, MD 20850, USA
COMMENT
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 85781: contig of 85781 bp in length
* 85782 85806: gap of unknown length
* 85807 102956: contig of 17150 bp in length
* 102957 102981: gap of unknown length
* 102982 105023: contig of 2042 bp in length.

FEATURES
source
1..105023
/organism="Trypanosoma brucei"
/mol_type="genomic DNA"
/isolate="GUTat10.1"
/db_xref="taxon:5691"
/chromosome="v"
/clone="RPC193-45E22"

ORIGIN
Query Match 99.4%; Score 9943.6; DB 2; Length 105023;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 9987; Conservative 0; Mismatches 9; Indels 8; Gaps 3;
QY 1 CCATCATATTTCTTATTTTGGCGGAGAGGAGACTTCTCTGTGCCAGGCTGG 60
Db 4284 CCATCATATTTCTTATTTTGGCGGAGAGGAGACTTCTCTGTGCCAGGCTGG 4343

QY 61 ACCAGTGGTGGGATCTTGGCTCACTGCAACCTCCACTCTCTGGGTCAAGTGAATCCCAA 120
DB 4344 ACCAGTGGTGGGATCTTGGCTCACTGCAACCTCCACTCTCTGGGTCAAGTGAATCCCAA 4403
QY 121 ATAGCTGGGATACAGGTGTATTAACCAATGCCAGCTAATTTCTGTATTTTAGCAGAT 180
DB 4404 ATAGCTGGGATACAGGTGTATTAACCAATGCCAGCTAATTTCTGTATTTTAGCAGAT 4463
QY 181 AAGGGGTTCACCATGTGGCCAGGCTGGTCTCCAACTCTCTGGCCCTCATGTGATCCACCC 240
DB 4464 AAGGGGTTCACCATGTGGCCAGGCTGGTCTCCAACTCTCTGGCCCTCATGTGATCCACCC 4523
QY 241 ACTTCGGCTTCCCAAGCAATGGGAGTATAGGTGTAGGACCACTATACCCGCTCTCACATC 300
DB 4524 ACTTCGGCTTCCCAAGCAATGGGAGTATAGGTGTAGGACCACTATACCCGCTCTCACATC 4583
QY 301 ATATTTCTAATCCCGAGACTGTAGAGCTGGTCTCTTTTCTAAGGATGTCAGTAGAG 360
DB 4584 ATATTTCTAATCCCGAGACTGTAGAGCTGGTCTCTTTTCTAAGGATGTCAGTAGAG 4643
QY 361 AAGTGGAGTTCGCCAAATTTACAGTTTTCACGTATTAGTCAAGTTTCTAATAATACAGTAAT 420
DB 4644 AAGTGGAGTTCGCCAAATTTACAGTTTTCACGTATTAGTCAAGTTTCTAATAATACAGTAAT 4703
QY 421 AATGTGAGAGCTGACATAGGAGCTAACTT - GGTTTTTTTTTTTTTTTTTTTTTTTTTCAAA 479
DB 4704 AATGTGAGAGCTGACATAGGAGCTAACTTAGTTTTTTTTTTTTTTTTTTTTTTTTTCAAA 4763
QY 480 TTCTCACTCAACTTTGATTTTGTCTAATAAGGACATTAAAGAAACCAAAACCACTCCA 539
DB 4764 TTCTCACTCAACTTTGATTTTGTCTAATAAGGACATTAAAGAAACCAAAACCACTCCA 4823
QY 540 CTATGGCTTATGGCACTATTTGATTTTTTAAATAAAGCGTATTATTAGCATCTAAAG 599
DB 4824 CTATGGCTTATGGCACTATTTGATTTTTTAAATAAAGCGTATTATTAGCATCTAAAG 4883
QY 600 TAGGAAGGACCTCAATAAATAGTCTTTGTTCTGGCCAGGAAACAGCGTTGTGACA 659
DB 4884 TAGGAAGGACCTCAATAAATAGTCTTTGTTCTGGCCAGGAAACAGCGTTGTGACA 4943
QY 660 ATTTGATTAACCTGTTTTCTAGGGTATGCTGTATTCTAGTTTAAACCTTGGCTGGGACG 719
DB 4944 ATTTGATTAACCTGTTTTCTAGGGTATGCTGTATTCTAGTTTAAACCTTGGCTGGGACG 5003
QY 720 CTAGCATTCAGTAAATACCTTTGTAATAAGCAAAATGAAACTTCTATGTATAGAA 779
DB 5004 CTAGCATTCAGTAAATACCTTTGTAATAAGCAAAATGAAACTTCTATGTATAGAA 5063
QY 780 ACTTAAGTCATTCACATTCATGTAGCAGAGTAATTTGAATATTTCTTTTCAATGTGTAGC 839
DB 5064 ACTTAAGTCATTCACATTCATGTAGCAGAGTAATTTGAATATTTCTTTTCAATGTGTAGC 5123
QY 840 TCTATCCCAGAACCAACAGAAATTTGGAATCTGAAAGGCCATCTTATAGTTTAAACCACT 899
DB 5124 TCTATCCCAGAACCAACAGAAATTTGGAATCTGAAAGGCCATCTTATAGTTTAAACCACT 5183
QY 900 GCGTTAAATPAGATAATAGAAAGATGTGTAATGTGGCAGTGACAACTTGAAGGTGTGACT 959
DB 5184 GCGTTAAATPAGATAATAGAAAGATGTGTAATGTGGCAGTGACAACTTGAAGGTGTGACT 5243
QY 960 AGAATCTGGGTCTCTGGAGTGTCTATTATATACACCAAGCTGGTCAACGCCCATGTG 1019
DB 5244 AGAATCTGGGTCTCTGGAGTGTCTATTATATACACCAAGCTGGTCAACGCCCATGTG 5303
QY 1020 TTGATCTCCATGTGTATAGCAACAAAGAAAGACTTCAGGACATTCCTTCTTTACCCT 1079
DB 5304 TTGATCTCCATGTGTATAGCAACAAAGAAAGACTTCAGGACATTCCTTCTTTACCCT 5363
QY 1080 AATCTCTGTATCTGCACTCTTATTTAGAAAAGCTTAAATGTTAAGATCTAGTTTATTCAAA 1139
DB 5364 AATCTCTGTATCTGCACTCTTATTTAGAAAAGCTTAAATGTTAAGATCTAGTTTATTCAAA 5423
QY 1140 ACTAAAGATAACAAGGAGTATGAGAAATTTCTATTTTCGGAGTGTAAAGGAGGAGATGTTTC 1199

DB 5424 ACTAAAGATAACAAGGAGTATGAGAAATTTCTATTTTCGGAGTGTAAAGGAGGAGATGTTTC 5483
QY 1200 CTTGGCTTCTCTGAGCCTCGAGCCTTCTTGGCTCTTCTTAAAGAAAGTAGAGAGGAGGAGGA 1259
DB 5484 CTTGGCTTCTCTGAGCCTCGAGCCTTCTTGGCTCTTCTTAAAGAAAGTAGAGAGGAGGAGGA 5543
QY 1260 AAGTAAAGTATGCTTTTGTGTTTTTAAAGGTATCTTTCTCTGGAGTAGTTTTCATGCTCCCTTTT 1319
DB 5544 AAGTAAAGTATGCTTTTGTGTTTTTAAAGGTATCTTTCTCTGGAGTAGTTTTCATGCTCCCTTTT 5603
QY 1320 GGTTTTCTTGGGTGGAAATTAAGTCTTAAGTTTTTAAAGTAGTTTGGGACTATTTTAAAAAACA 1379
DB 5604 GGTTTTCTTGGGTGGAAATTAAGTCTTAAGTTTTTAAAGTAGTTTGGGACTATTTTAAAAAACA 5663
QY 1380 ATGCTTATCCAAATGTTTGGCATATAGGCGAGAGGTATTTGGCTTTTAAAGTTAATTTCTTCT 1439
DB 5664 ATGCTTATCCAAATGTTTGGCATATAGGCGAGAGGTATTTGGCTTTTAAAGTTAATTTCTTCT 5723
QY 1440 CCAGGAGTGAATAATTAGCTTTCTAAACCAAGCAGCAGAGCTAAATTAAGTAAATTTTCCA 1499
DB 5724 CCAGGAGTGAATAATTAGCTTTCTAAACCAAGCAGCAGAGCTAAATTAAGTAAATTTTCCA 5783
QY 1500 CTTGGCAGTGCATGATGTGAAGGTAGATTAAAAAATGAGAGGCCCAATTTTCTGATG 1559
DB 5784 CTTGGCAGTGCATGATGTGAAGGTAGATTAAAAAATGAGAGGCCCAATTTTCTGATG 5843
QY 1560 AAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATACATTCA 1619
DB 5844 AAGACTAAGCCATGTTGAAACAGCCCTGTTGAGGATTTTATTTTAAATCTATACATTCA 5903
QY 1620 CAAAGGAGCTTTGTGTATGTTTCCCTATTTTGGCTAGGAAAGCCCAACCCAG 1679
DB 5904 CAAAGGAGCTTTGTGTATGTTTCCCTATTTTGGCTAGGAAAGCCCAACCCAG 5963
QY 1680 TGCCTTTGAAAGCGCAAAAGTCTGTTGAAAGCAAGCTGGGATTTGAAACAGTGGATTCAGGT 1739
DB 5964 TGCCTTTGAAAGCGCAAAAGTCTGTTGAAAGCAAGCTGGGATTTGAAACAGTGGATTCAGGT 6023
QY 1740 TTCGATATCCAGTCAACCAAAATATATCAGGGTTCCTCTGCCCAGATGAGTGACATTT 1799
DB 6024 TTCGATATCCAGTCAACCAAAATATATCAGGGTTCCTCTGCCCAGATGAGTGACATTT 6083
QY 1800 CTGAGTGTAAAGTATTTCTGAAATGGGATTTTAAAGAAAGTTTCTGTATTTCTGTGCT 1859
DB 6084 CTGAGTGTAAAGTATTTCTGAAATGGGATTTTAAAGAAAGTTTCTGTATTTCTGTGCT 6143
QY 1860 CAATTTGTGACCTCTGTATGTGCAAAATCTCTAAAGGGGTGTTTGGGCACTTAGATTTTC 1919
DB 6144 CAATTTGTGACCTCTGTATGTGCAAAATCTCTAAAGGGGTGTTTGGGCACTTAGATTTTC 6203
QY 1920 TTGGATGCGAGATTTGTTGTATATGAAACAAATTTTAAATTTGTTTGTATACACTGGATT 1979
DB 6204 TTGGATGCGAGATTTGTTGTATATGAAACAAATTTTAAATTTGTTTGTATACACTGGATT 6263
QY 1980 TAAAAATAGTTTAAAGTGTGTTTTTAAATTTTCTTAAATTTTCACTTAAATTTTCACTAGTC 2039
DB 6264 TAAAAATAGTTTAAAGTGTGTTTTTAAATTTTCTTAAATTTTCACTTAAATTTTCACTAGTC 6323
QY 2040 TTTAGATTTAGGAGGCTGTTGATGGCATCCACATGTGCAATTTTATAGTGGCAATTTAAAAATG 2099
DB 6324 TTTAGATTTAGGAGGCTGTTGATGGCATCCACATGTGCAATTTTATAGTGGCAATTTAAAAATG 6383
QY 2100 TATTTCAGCTGAATTTTAAACAAATTTCTGACCTTAAACCTTGAATTTTATAGTTAAGTCGGTA 2159
DB 6384 TATTTCAGCTGAATTTTAAACAAATTTCTGACCTTAAACCTTGAATTTTATAGTTAAGTCGGTA 6443
QY 2160 AAGCACTGATTTTAAACTGGATTTTAACTGGATGAAATTTCTGATTTTAAAGTACTGAC 2219
DB 6444 AAGCACTGATTTTAAACTGGATTTTAACTGGATGAAATTTCTGATTTTAAAGTACTGAC 6503
QY 2220 TGGATAAAATGCCAAATGATTTTAAATTAACAAGCAGCTTTTAAACAGGATGCCCTATATATTAG 2279

Db 6504 TGGATAAATGCCAATGATTTAAATTAACAAGCACGTTTAAACAGGATGCCCTATATATTAG 6563
Qy
2280 TTAAAGTGAAGCAATTTGAATTTAGGTACCTTCTCTGCTGCGTGGAAAAAGACCGGTATGACT 2339
Db 6564 TTAAAGTGAAGCAATTTGAATTTAGGTACCTTCTCTGCTGCGTGGAAAAAGACCGGTATGACT 6623
Qy 2340 CACCCACACCGCCTTCTCTGCTGCTGAGTGTAGCTTAACCGTCTCTGTTTTTTTCCCTC 2399
Db 6624 CACCCACACCGCCTTCTCTGCTGCTGAGTGTAGCTTAACCGTCTCTGTTTTTTTCCCTC 6683
Qy 2400 TAGGGTTTGGAAATCCCTTGTCTCCAGGTTGTGCGAATTTGACTTCTTGTCAATTTGAATC 2459
Db 6684 TAGGGTTTGGAAATCCCTTGTCTCCAGGTTGTGCGAATTTGACTTCTTGTCAATTTGAATC 6743
Qy 2460 ACTCAATTTCAATGGAGACAAAGAGAACTAATGCTTTTGTGCTGATTCATATTTGAATTCGAGG 2519
Db 6744 ACTCAATTTCAATGGAGACAAAGAGAACTAATGCTTTTGTGCTGATTCATATTTGAATTCGAGG 6803
Qy 2520 CATTGGGACCCCTGTATGCCCTTGTGTTGGAAAGAACCAAGTGACACCACTCACTGAGCTTC 2579
Db 6804 CATTGGGACCCCTGTATGCCCTTGTGTTGGAAAGAACCAAGTGACACCACTCACTGAGCTTC 6863
Qy 2580 CTAAAGTTTCGAAGAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATATAAATAT 2639
Db 6864 CTAAAGTTTCGAAGAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATATAAATAT 6923
Qy 2640 TTGCTCTGTTTTTGGAAACCCAGGCTGTTAGAGGGTGAGTGA CAAGTCTTTACAAGTGG 2699
Db 6924 TTGCTCTGTTTTTGGAAACCCAGGCTGTTAGAGGGTGAGTGA CAAGTCTTTACAAGTGG 6983
Qy 2700 CTTTATTCCACTCCAGAAATGCCCCACCGAACTTTGAGATTATATGCAATCGAAAGTG 2759
Db 6984 CTTTATTCCACTCCAGAAATGCCCCACCGAACTTTGAGATTATATGCAATCGAAAGTG 7043
Qy 2760 ACAGAAATCATGCCAACTCAATCCCTCTTAATGTATCATGGATGGGCCAGAACTGATGGC 2819
Db 7044 ACAGAAATCATGCCAACTCAATCCCTCTTAATGTATCATGGATGGGCCAGAACTGATGGC 7103
Qy 2820 AGCTCTCTTGGCAGTCCGATGAGATGAGAGATGCTTGTCTAATGAAAGGGACCGCTGTT 2879
Db 7104 AGCTCTCTTGGCAGTCCGATGAGATGAGAGATGCTTGTCTAATGAAAGGGACCGCTGTT 7163
Qy 2880 GTTTCATTTCCGAGCTACACAAGAAAAAATGTCAATCAATCGAGGGGTATATGCCCTTG 2939
Db 7164 GTTTCATTTCCGAGCTACACAAGAAAAAATGTCAATCAATCGAGGGGTATATGCCCTTG 7223
Qy 2940 GATTGCTATTTCTGCAGCCAGACCTTTACACATTCAGAAAGCTTTAATAAATCATGTCTTTA 2999
Db 7224 GATTGCTATTTCTGCAGCCAGACCTTTACACATTCAGAAAGCTTTAATAAATCATGTCTTTA 7283
Qy 3000 ATGCAACACCGGCCTTACCTCTGTGAAACAGCAGTTCCTTGGGTTGAAGCAGATATCTC 3059
Db 7284 ATGCAACACCGGCCTTACCTCTGTGAAACAGCAGTTCCTTGGGTTGAAGCAGATATCTC 7343
Qy 3060 AGTCCGCTTGATAAAGTCAAGTGGCAAGACAGACCTCCCAAGGAAAAAGTTGCAAGGAA 3119
Db 7344 AGTCCGCTTGATAAAGTCAAGTGGCAAGACAGACCTCCCAAGGAAAAAGTTGCAAGGAA 7403
Qy 3120 AATGAATTTAGCTGTGAGGTATGTGGCAGACATTTAGAGTCCGCTTTTGTATGTTGAGATC 3179
Db 7404 AATGAATTTAGCTGTGAGGTATGTGGCAGACATTTAGAGTCCGCTTTTGTATGTTGAGATC 7463
Qy 3180 CACATGAGAACACAAAGATTTCTTCACTTACGGGTGTAAACATGTGCGGAAAGATTC 3239
Db 7464 CACATGAGAACACAAAGATTTCTTCACTTACGGGTGTAAACATGTGCGGAAAGATTC 7523
Qy 3240 AAGGAGCCTTGTGTTCTTAAATATCACATGCGGACACATAATGGCAAAATCGGGGCCAGA 3299
Db 7524 AAGGAGCCTTGTGTTCTTAAATATCACATGCGGACACATAATGGCAAAATCGGGGCCAGA 7583
Qy 3300 AGCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCAGGTG 3359
Db 7584 AGCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCGTCAGGTG 7643

Qy 3360 CACGCGCGGAGAGATCTCTCTCTCTTACAAAATCTGCATGGTTTGTGGCTTCTCTATTTT 3419
Db 7644 CACGCGCGGAGAGATCTCTCTCTCTTACAAAATCTGCATGGTTTGTGGCTTCTCTATTTT 7703
Qy 3420 CCAATAAAGAAAAGTCTAATTTAGCACCGCAAGGTGCACACAAAATAACTCTCTTTCCGT 3479
Db 7704 CCAATAAAGAAAAGTCTAATTTAGCACCGCAAGGTGCACACAAAATAACTCTCTTTCCGT 7763
Qy 3480 ACCAGCAGCGCGCAGACAGACTCTCCACAAGSAGGAATGCCGTCTCGAGGGAGGACTTC 3539
Db 7764 ACCAGCAGCGCGCAGACAGACTCTCCACAAGSAGGAATGCCGTCTCGAGGGAGGACTTC 7823
Qy 3540 CTGCAAGTTGTTCAACTTGAGACCAAAAATCTCACTCTGAAAACGGGGAAGAGCTGTCTCAGA 3599
Db 7824 CTGCAAGTTGTTCAACTTGAGACCAAAAATCTCACTCTGAAAACGGGGAAGAGCTGTCTCAGA 7883
Qy 3600 TGCATCCCTCAGCTCGATCCGTTTACCACTCTTCAGGCTTGGCAGCTGGCTACCAAGGA 3659
Db 7884 TGCATCCCTCAGCTCGATCCGTTTACCACTCTTCAGGCTTGGCAGCTGGCTACCAAGGA 7943
Qy 3660 AAAGTTGCCATTTGCAAGAAAGTGAAGAAATCGGGGCAAGAAAGGAGCAACCAACGAC 3719
Db 7944 AAAGTTGCCATTTGCAAGAAAGTGAAGAAATCGGGGCAAGAAAGGAGCAACCAACGAC 8003
Qy 3720 GATTGAGTTCGAGAACGAGCTTGGAGAAACAAATAAGGCAGTGTGTCAGGCTCTCG 3779
Db 8004 GATTGAGTTCGAGAACGAGCTTGGAGAAACAAATAAGGCAGTGTGTCAGGCTCTCG 8063
Qy 3780 CAAAGAAAGAGAGTGCACACACTCCACGGCGAAGGCCCTCCGTGAGCGGATCCC 3839
Db 8064 CAAAGAAAGAGAGTGCACACACTCCACGGCGAAGGCCCTCCGTGAGCGGATCCC 8123
Qy 3840 AAGTTACCCAGTAGCAAGSAGAACCCACTCACTCTCGAGTGGCGGCAAGCTTTTCAGA 3899
Db 8124 AAGTTACCCAGTAGCAAGSAGAACCCACTCACTCTCGAGTGGCGGCAAGCTTTTCAGA 8183
Qy 3900 ACCTACCAACAGCTGCTTGCACCTCAGGCTTCCACAGAGGACCGGAGGCGCGCG 3959
Db 8184 ACCTACCAACAGCTGCTTGCACCTCAGGCTTCCACAGAGGACCGGAGGCGCGCG 8243
Qy 3960 GAGTCCGCCACCATGCTGTGGAACGGGAGGACCGCGGACCGTGTCTCTGACCTCGCC 4019
Db 8244 GAGTCCGCCACCATGCTGTGGAACGGGAGGACCGCGGACCGTGTCTCTGACCTCGCC 8303
Qy 4020 GCCCTCTTGGATGAAAAATGGAGCCGTGGAATCGAGGGGAAGTGGTTCGGAAGACGATCT 4079
Db 8304 GCCCTCTTGGATGAAAAATGGAGCCGTGGAATCGAGGGGAAGTGGTTCGGAAGACGATCT 8363
Qy 4080 GAGGATGGCTTCCGAAAGGAATCCATCTGGGTAGCTGCCGTGCTCCGTCCCGTGTG 4139
Db 8364 GAGGATGGCTTCCGAAAGGAATCCATCTGGGTAGCTGCCGTGCTCCGTCCCGTGTG 8423
Qy 4140 TTCCGCCCTGTGCTGTCTCTCCCGCTCTCCCGCTCTCTATTCCCATCTCCAGACAAAG 4199
Db 8424 TTCCGCCCTG----TGTCTGTCTCCCGCTCTCCCGCTCTCTATTCCCATCTCCAGACAAAG 8479
Qy 4200 TTGCGCAAGGAATGGGGTTTGGAGACCGCAGAGTCAAGTCCAGGCTCTTTTTTGGTATCACCTC 4259
Db 8480 TTGCGCAAGGAATGGGGTTTGGAGACCGCAGAGTCAAGTCCAGGCTCTTTTTTGGTATCACCTC 8539
Qy 4260 TGTGTAAGTCAATTTAACTCTCAGGGCCTTAATTTTTCTCATTTCTGTAATAACAGGGTGG 4319
Db 8540 TGTGTAAGTCAATTTAACTCTCAGGGCCTTAATTTTTCTCATTTCTGTAATAACAGGGTGG 8599
Qy 4320 AGTTAAGAGTCTCTCTGTTCTGAAAAATATATATATTTTTTAAACGTGTATCGTTTTG 4379
Db 8600 AGTTAAGAGTCTCTCTGTTCTGAAAAATATATATATTTTTTAAACGTGTATCGTTTTG 8659
Qy 4380 CTCACAAAACACACTTTTAAAAAATAAATTGTTGATCCAGGCCAAATGCACTGCTTC 4439
Db 8660 CTCACAAAACACACTTTTAAAAAATAAATTGTTGATCCAGGCCAAATGCACTGCTTC 8719

QY	4440	TTAACTGGGCGAATTTTGTCCCAATCAGTATCTGGCAATGTCTGGAGGCAATTTGGTTG	4499
Db	8720	TTAACTGGGCGAATTTTGTCCCAATCAGTATCTGGCAATGTCTGGAGGCAATTTGGTTG	8779
QY	4500	TCATACCTGCTGTGTGGGTGCTGTGGCATCCAGTGGCGCAGAGCCAGGACACTGC	4559
Db	8780	TCATACCTGCTGTGTGGGTGCTGTGGCATCCAGTGGCGCAGAGCCAGGACACTGC	8839
QY	4560	TCAGCATGGTACAGTGCACAGGACAGCCCATCATCAAGAAATTAATCTGGTCCCAATGT	4619
Db	8840	TCAGCATGGTACAGTGCACAGGACAGCCCATCATCAAGAAATTAATCTGGTCCCAATGT	8899
QY	4620	CAATAGTTTGAGCAATGAGAGACCTTAGCCTTCACTTAAAGTTTTCCTGGCGTTCTCGATC	4679
Db	8900	CAATAGTTTGAGCAATGAGAGACCTTAGCCTTCACTTAAAGTTTTCCTGGCGTTCTCGATC	8959
QY	4680	TTTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGAG---TGATCAACTAGAG	4736
Db	8960	TTTTTCTGTAGTGAATTTCTAGTGGCCATAAAAGGTACTGGAG---TGATCAACTAGAG	9019
QY	4737	CCAGGAATATATTTGGGCGAGCGTTTGGTGTCTGTCCAAACCTTGTCTCTGTCTGG	4796
Db	9020	CCAGTAATATATTTGGGCGAGCGTTTGGTGTCTGTCCAAACCTTGTCTCTGTCTGG	9079
QY	4797	CAAGCTAGTATCCATTTATAGTACCTCAGGAACCCAAATGAATTTGTCAATAAATACAAG	4856
Db	9080	CAAGCTAGTATCCATTTATAGTACCTCAGGAACCCAAATGAATTTGTCAATAAATACAAG	9139
QY	4857	GAATGTGAGCACACTGAAGCAATTTTAAAGAGGCTCATTTGCTCAGAGCAATTTTCAGT	4916
Db	9140	GAATGTGAGCACACTGAAGCAATTTTAAAGAGGCTCATTTGCTCAGAGCAATTTTCAGT	9199
QY	4917	GTACTAGTGCAATTTATAGAAAGAGAGGTGATCACTGAAGGCAATGCTCACATAATATTC	4976
Db	9200	GTACTAGTGCAATTTATAGAAAGAGAGGTGATCACTGAAGGCAATGCTCACATAATATTC	9259
QY	4977	CTGAGCCCTGGTGGCGTTATCTAGGCAAGGATTCACCTGTGTGGAGTTGGCCCC	5036
Db	9260	CTGAGCCCTGGTGGCGTTATCTAGGCAAGGATTCACCTGTGTGGAGTTGGCCCC	9319
QY	5037	ATCCTCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTGGATAGAGGAT	5096
Db	9320	ATCCTCACTGTAGCCAGAGCTTCTCCTATCAGAGTTTAGTATTTTGTGGATAGAGGAT	9379
QY	5097	CTTGCTGCTTAAACAGTTGAAAGACCCCTGATGGCGAGCCGTAAATGACAGCGCAATG	5156
Db	9380	CTTGCTGCTTAAACAGTTGAAAGACCCCTGATGGCGAGCCGTAAATGACAGCGCAATG	9439
QY	5157	ATGGGAACATGAATCCGGTCTTAGGGAAGCATCTGTCAAAAGTGGTCCCTGGTTAAACAAAG	5216
Db	9440	ATGGGAACATGAATCCGGTCTTAGGGAAGCATCTGTCAAAAGTGGTCCCTGGTTAAACAAAG	9499
QY	5217	TGCCTCCTCTCCTCAGTGCACTTGATTTGTGTGCTTGAATTTCTTGGGAAACCTGGGTGTA	5276
Db	9500	TGCCTCCTCCTCCTCAGTGCACTTGATTTGTGTGCTTGAATTTCTTGGGAAACCTGGGTGTA	9559
QY	5277	TGAGACCCACGATGAATTTGCCACACAGATTGATTCGACTTTCCTTCCCTGCTCTTCA	5336
Db	9560	TGAGACCCACGATGAATTTGCCACACAGATTGATTCGACTTTCCTTCCCTGCTCTTCA	9619
QY	5337	GCAGTGCCAGTTTCTTTTCTGATCATGTGATTTGACGTGAGAACTGTAGTCTGTATATCA	5396
Db	9620	GCAGTGCCAGTTTCTTTTCTGATCATGTGATTTGACGTGAGAACTGTAGTCTGTATATCA	9679
QY	5397	AATCTTTAGAAATTTTGTAGTTTCTGGGACACAGGAAACCCAGCACTTAGCATACTAC	5456
Db	9680	AATCTTTAGAAATTTTGTAGTTTCTGGGACACAGGAAACCCAGCACTTAGCATACTAC	9739
QY	5457	AAATCTAATGTCTTAATGCAATCAAAAGAGGCTTTTAAACACAGACTCCAGTTAGCTA	5516
Db	9740	AAATCTAATGTCTTAATGCAATCAAAAGAGGCTTTTAAACACAGACTCCAGTTAGCTA	9799
QY	5517	AGTGGTTTCTGTAGTGGCGGTACTGTTGAGGGGCCCTGTGAGATGCCCCAGTTCCCTG	5576

Db	9800	AGTGGTTTCTGTAGTGGCGGTACTGTTGAGGGGCCCTGTGAGATGCCCCAGTTCCCTG	9859
QY	5577	AAAGAAATGAAGAGCCAGTTACCGGTAGTGTGTGGGAAACCATGGGCTAGATCATCAG	5636
Db	9860	AAAGAAATGAAGAGCCAGTTACCGGTAGTGTGTGGGAAACCATGGGCTAGATCATCAG	9919
QY	5637	GCAGGACAGAAATCCCTGGCTGTGGGTGGGAGACCCAGCTTGGCGTTGAGTTCTGGTTC	5696
Db	9920	GCAGGACAGAAATCCCTGGCTGTGGGTGGGAGACCCAGCTTGGCGTTGAGTTCTGGTTC	9979
QY	5697	TACCACTGGCTGTGTGTGTGACCAATTAATGAGTTGCTTAAACCTTTCTTGTCTACTATTC	5756
Db	9980	TACCACTGGCTGTGTGTGTGACCAATTAATGAGTTGCTTAAACCTTTCTTGTCTACTATTC	10039
QY	5757	CCTGTTTGCAAAATGTTCAATGACCCCTGTCTTCACCTCCCAAGGACAAATTTCAACAG	5816
Db	10040	CCTGTTTGCAAAATGTTCAATGACCCCTGTCTTCACCTCCCAAGGACAAATTTCAACAG	10099
QY	5817	CCTATTGTAAAGAGATCACAGTCCCTTAAAGAAATATAACTGTAAAGTCAGAGGTGATGC	5876
Db	10100	CCTATTGTAAAGAGATCACAGTCCCTTAAAGAAATATAACTGTAAAGTCAGAGGTGATGC	10159
QY	5877	TTGAAAGAGCAGGAAACCCAGGTAGATGTGAAATGTCTATGTCTTCTTAAAGAAAGG	5936
Db	10160	TTGAAAGAGCAGGAAACCCAGGTAGATGTGAAATGTCTATGTCTTCTTAAAGAAAGG	10219
QY	5937	CAITTCATAGCTTTTGGATATGACCAACATACCAATAAATCTCTGACACATAGTTGGGAG	5996
Db	10220	CAITTCATAGCTTTTGGATATGACCAACATACCAATAAATCTCTGACACATAGTTGGGAG	10279
QY	5997	TGGGAAATTTGCAACACGCCCTGATTAACCCAGCTAGTTTGGGTATGATTTCTAGAGAA	6056
Db	10280	TGGGAAATTTGCAACACGCCCTGATTAACCCAGCTAGTTTGGGTATGATTTCTAGAGAA	10339
QY	6057	AAAAAGCTGGCCATTTCTGTATTGGGAAATGATTTTCTTAAACTTATATTTCTTAGTA	6116
Db	10340	AAAAAGCTGGCCATTTCTGTATTGGGAAATGATTTTCTTAAACTTATATTTCTTAGTA	10399
QY	6117	GTCTAGATTTATCATATTGTACTATCATCTGGCTTTTAAAGACTTAAAGAGTCAAGT	6176
Db	10400	GTCTAGATTTATCATATTGTACTATCATCTGGCTTTTAAAGACTTAAAGAGTCAAGT	10459
QY	6177	AAATTTTCTTCTTTTAGACACTATATAGATCATCAAGGCTCTGTCTTACAGT	6236
Db	10460	AAATTTTCTTCTTTTAGACACTATATAGATCATCAAGGCTCTGTCTTACAGT	10519
QY	6237	GGATAGTGATATGATCTACAGTGAGGGGACATTTATTTAAACCTTAAACATTCATGTGT	6296
Db	10520	GGATAGTGATATGATCTACAGTGAGGGGACATTTATTTAAACCTTAAACATTCATGTGT	10579
QY	6297	TTGGGGGTGGTATTTTAAAGGAGACCTCTGATTTGTCTTTTGGAGGGCTGGTGTGTGT	6356
Db	10580	TTGGGGGTGGTATTTTAAAGGAGACCTCTGATTTGTCTTTTGGAGGGCTGGTGTGTGT	10639
QY	6357	TGAAGTCTGCTCCTTCCAGTGAGCTCTAATCTCTGATGACAGTGGAGACATTCG	6416
Db	10640	TGAAGTCTGCTCCTTCCAGTGAGCTCTAATCTCTGATGACAGTGGAGACATTCG	10699
QY	6417	TCCTATTGTCTTCGAGAACTAAAGCCAAACACTGTCTCTGGGACAGGTTTTCATTTG	6476
Db	10700	TCCTATTGTCTTCGAGAACTAAAGCCAAACACTGTCTCTTGGGACAGGTTTTCATTTG	10759
QY	6477	TCAGATCTCTTTTCGGCCACATGAGTGTGTTGTGGCAATACAGCTCTCTTCCAAAATTT	6536
Db	10760	TCAGATCTCTTTTCGGCCACATGAGTGTGTTGTGGCAATACAGCTCTCTTCCAAAATTT	10819
QY	6537	GCTAAATTTTGACAGACTTTCCTAGTGTCTGCCAAATGCCAGACTTTCTTTCTGTTCA	6596
Db	10820	GCTAAATTTTGACAGACTTTCCTAGTGTCTGCCAAATGCCAGACTTTCTTTCTGTTCA	10879
QY	6597	AGATTAAGTTGTGCTTGTCTGCTCTAGTGTGTCAGTTGTTTAAATCTTAACCTTAAACGCG	6656

Db	10880	AGATTAAAGTTGCTTGCTGCCCTCTAGTGGTCAGTTGTTTAAATCCTTAACCTTAAACGGC	10939
Qy	6657	TTATTTTTCCTCGTGGTGGGAAGTTGACGGTTGTAAATGGCTCATTTTCTAAATTT	6716
Db	10940	TTATTTTTCCTCGTGGTGGGAAGTTGACGGTTGTAAATGGCTCATTTTCTAAATTT	10999
Qy	6717	ATTCTGMAAGAATAAATTTTCCGCCAGTAGTATGTCACCTTCAGTTTGGCCAGATCC	6776
Db	11000	ATTCTGMAAGAATAAATTTTCCGCCAGTAGTATGTCACCTTCAGTTTGGCCAGATCC	11059
Qy	6777	TGCTGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGCAATTCAGTCTATGAAATGAT	6836
Db	11060	TGCTGCTCAGAGACACTGAGAACCGGAAGCTGCCCGGCAATTCAGTCTATGAAATGAT	11119
Qy	6837	CTTCTGTGTATTAAGGCAAAACGAAGAACTGAAATGTTTAATAGTGTACTCTGCTGATCCC	6896
Db	11120	CTTCTGTGTATTAAGGCAAAACGAAGAACTGAAATGTTTAATAGTGTACTCTGCTGATCCC	11179
Qy	6897	AGAAAAACAAAAAANAATCATGTTATAACAACCTCTAAACCTTCAACAAACCTCCAACAG	6956
Db	11180	AGAAAAACAAAAAANAATCATGTTATAACAACCTCTAAACCTTCAACAAACCTCCAACAG	11239
Qy	6957	CATTGGTGTGTCTAGCCGTTTTGTTCTAAACCGGATGTTATATAAAGAAATTTTTTCA	7016
Db	11240	CATTGGTGTGTCTAGCCGTTTTGTTCTAAACCGGATGTTATATAAAGAAATTTTTTCA	11299
Qy	7017	TGCTTTCAAAAATGTTTATGTCAGAAATATTTAAAGTCAGCATGCTTATTCAGGTACTTT	7076
Db	11300	TGCTTTCAAAAATGTTTATGTCAGAAATATTTAAAGTCAGCATGCTTATTCAGGTACTTT	11359
Qy	7077	CAGCTACCTTCTTATATATAAATTTTGTCTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT	7136
Db	11360	CAGCTACCTTCTTATATAAATTTTGTCTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT	11419
Qy	7137	AAAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTTGTAAGAGTAAAGTAAAGT	7196
Db	11420	AAAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTTGTAAGAGTAAAGTAAAGT	11479
Qy	7197	AAATTTATACCTCAATATTCATCTCAGAACGGATACAGGTAAAGAACTTTTATTTTTTAA	7256
Db	11480	AAATTTATACCTCAATATTCATCTCAGAACGGATACAGGTAAAGAACTTTTATTTTTTAA	11539
Qy	7257	ACCATGCATTTAGTAAATATGATGATTTATTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTT	7316
Db	11540	ACCATGCATTTAGTAAATATGATGATTTATTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTT	11599
Qy	7317	CTGCCAGATCCCTGGACTAGCTTAAAGGATAAATATGTAGCATGTTGATTCAGTGGTTAT	7376
Db	11600	CTGCCAGATCCCTGGACTAGCTTAAAGGATAAATATGTAGCATGTTGATTCAGTGGTTAT	11659
Qy	7377	TTTTTATCTTTTGTAGTGCCATTTGTAATCTGAGGCAATGTTTCTTATTTGAGTTTCTTTCT	7436
Db	11660	TTTTTATCTTTTGTAGTGCCATTTGTAATCTGAGGCAATGTTTCTTATTTGAGTTTCTTTCT	11719
Qy	7437	TTCTTTCTTTTGTGTTTGTAGACGGAGTCTTGTCTGTACCTCCGCTGAGTGGCAGT	7496
Db	11720	TTCTTTCTTTTGTGTTTGTAGACGGAGTCTTGTCTGTACCTCCGCTGAGTGGCAGT	11779
Qy	7497	GGTGCAATTTCCGGCTCACCTGCAGCTTCAACCTCCCTGGTTCAAGCAATACTCCTCGCTCA	7556
Db	11780	GGTGCAATTTCCGGCTCACCTGCAGCTTCAACCTCCCTGGTTCAAGCAATACTCCTCGCTCA	11839
Qy	7557	GCCTCCCAAGTAGTTGGGATTAACAGGTACCTGCCACCAACCCCGGCTAATTTCTGTATTT	7616
Db	11840	GCCTCCCAAGTAGTTGGGATTAACAGGTACCTGCCACCAACCCCGGCTAATTTCTGTATTT	11899
Qy	7617	TTAGTAGAGATGGGTTTCAACATGCTGGCCAGGCTGTTTTCGAACTCTGTACCTCAAGT	7676
Db	11900	TTAGTAGAGATGGGTTTCAACATGCTGGCCAGGCTGTTTTCGAACTCTGTACCTCAAGT	11959
Qy	7677	GATCCGCTCACCTTGGCCTCCCATAGTGTGGCCTCCCATAGTGTGGCCTCCCATAGTGTGGCCT	7736
Db	11960	GATCCGCTCACCTTGGCCTCCCATAGTGTGGCCTCCCATAGTGTGGCCTCCCATAGTGTGGCCT	12019
Qy	7737	GAGCACCGCGCCCGGACAAAAGTTCAATTTGTTAGTTTATGACTGCTATGCTGACTCT	7796
Db	12020	GAGCACCGCGCCCGGACAAAAGTTCAATTTGTTAGTTTATGACTGCTATGCTGACTCT	12079
Qy	7797	TATCTTATTAAGAGCTACAGTATTTTAAATGCTGCATCTTATGCTCTTTATGATGAGAA	7856
Db	12080	TATCTTATTAAGAGCTACAGTATTTTAAATGCTGCATCTTATGCTCTTTATGATGAGAA	12139
Qy	7857	TGAAATGAGAAATCTATTTTATGATGCTTGAGATTTGTAAAGGAGCTATGACATCATGAT	7916
Db	12140	TGAAATGAGAAATCTATTTTATGATGCTTGAGATTTGTAAAGGAGCTATGACATCATGAT	12199
Qy	7917	AGGAGCTGCGTAGATTTTGAATTTTCATCTTCCACTTACTATCTGTCGACCTTGGGC	7976
Db	12200	AGGAGCTGCGTAGATTTTGAATTTTCATCTTCCACTTACTATCTGTCGACCTTGGGC	12259
Qy	7977	AAGTTTATTTAACTTTTGTGCTTTTGTAGTTTCTTCTGCTGTAAGTAAAGTAAAGTAAAGT	8036
Db	12260	AAGTTTATTTAACTTTTGTGCTTTTGTAGTTTCTTCTGCTGTAAGTAAAGTAAAGTAAAGT	12319
Qy	8037	TTTCCCTAGGCTGTTAGGAAAGATTTAAATTAAGTTAAGTGTGCTGTTTAAATTTTCTAT	8096
Db	12320	TTTCCCTAGGCTGTTAGGAAAGATTTAAATTAAGTTAAGTGTGCTGTTTAAATTTTCTAT	12379
Qy	8097	TGAAGATAGGCATTCATATTTCAATATTTCAATATTTCAATATTTCAATATTTCAATATTT	8156
Db	12380	TGAAGATAGGCATTCATATTTCAATATTTCAATATTTCAATATTTCAATATTTCAATATTT	12439
Qy	8157	AGAAATCTTATGTAGTAGTAGATCGAGAAAGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGG	8216
Db	12440	AGAAATCTTATGTAGTAGTAGATCGAGAAAGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGG	12499
Qy	8217	ATAATAGATATTTGATCTATTTTCACTGCTTTTTCATACACTTTCTATAATAAAGTGCATTT	8276
Db	12500	ATAATAGATATTTGATCTATTTTCACTGCTTTTTCATACACTTTCTATAATAAAGTGCATTT	12559
Qy	8277	TCTTGCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGAGCCAGAA	8336
Db	12560	TCTTGCCTTAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGAGCCAGAA	12619
Qy	8337	GACATCTCTAGGTATCACTTGGAGAGACATCACAAAGGAAAAAACAACCGATGTTGCTGC	8396
Db	12620	GACATCTCTAGGTATCACTTGGAGAGACATCACAAAGGAAAAAACAACCGATGTTGCTGC	12679
Qy	8397	TGAAGTCAAGAACGATGTTAAATAATCAGGACACTGAAGATGCACTATTAAACCGCTGACAG	8456
Db	12680	TGAAGTCAAGAACGATGTTAAATAATCAGGACACTGAAGATGCACTATTAAACCGCTGACAG	12739
Qy	8457	TGCGCAAAACCAAAATTTTGAAGAATTTTGTGATGCTGCCAAGATGTTTACAGGAGTCC	8516
Db	12740	TGCGCAAAACCAAAATTTTGAAGAATTTTGTGATGCTGCCAAGATGTTTACAGGAGTCC	12799
Qy	8517	ACCTGCAAGCAGCTTAAAGGAGATGCTTCTGTTTTTTCAGAAATGTTTGGGAGCGCTGT	8576
Db	12800	ACCTGCAAGCAGCTTAAAGGAGATGCTTCTGTTTTTTCAGAAATGTTTGGGAGCGCTGT	12859
Qy	8577	CCTCTCACGACACAATAAGATATCTAGGATTTTCCATAAAAATGAGCTGATGACAGTGC	8636
Db	12860	CCTCTCACGACACAATAAGATATCTAGGATTTTCCATAAAAATGAGCTGATGACAGTGC	12919
Qy	8637	TGATTAAGTGAATAAAAAACCTTACCTGCTTACCTGAGCTGTTTAAAAAAGAGATCAGC	8696
Db	12920	TGATTAAGTGAATAAAAAACCTTACCTGCTTACCTGAGCTGTTTAAAAAAGAGATCAGC	12979
Qy	8697	AGTTGAAACTCAGGCAAAATTAACCTCATCTGTAGAAACCAAGCGGATGTTTCTCTCTCC	8756
Db	12980	AGTTGAAACTCAGGCAAAATTAACCTCATCTGTAGAAACCAAGCGGATGTTTCTCTCTCC	13039
Qy	8757	GGATGGCAGTACCCCAATAACCTTTGAAGTTAGTCCCAAGAGAGCAAAACGAGACCCG	8816
Db	13040	GGATGGCAGTACCCCAATAACCTTTGAAGTTAGTCCCAAGAGAGCAAAACGAGACCCG	13099

QY 8817 AGCTGACTGCAGATACAGGCAAGTGTGGATTTGTACGAAAAACCTTTAAATTTATCCGT 8876
DB 13100 AGCTGACTGCAGATACAGGCAAGTGTGGATTTGTACGAAAAACCTTTAAATTTATCCGT 13159
QY 8877 GGGGGCTTTTACAAATTTGCCCGCAATTTCTTTGAGTAAAGTTGATTCACAGTATCAC 8936
DB 13160 GGGGGCTTTTACAAATTTGCCCGCAATTTCTTTGAGTAAAGTTGATTCACAGTATCAC 13219
QY 8937 CTGTCCATTTTGTACCTTCAAGACATTTTATCCAGAAAGTTTAAATGATGACACAGAGACT 8996
DB 13220 CTGTCCATTTTGTACCTTCAAGACATTTTATCCAGAAAGTTTAAATGATGACACAGAGACT 13279
QY 8997 GGAGCATAAATACATCTGAGCTTCTATAAAAACTGTGCGAAACAAGTCTTGTCTTGAAG 9056
DB 13280 GGAGCATAAATACATCTGAGCTTCTATAAAAACTGTGCGAAACAAGTCTTGTCTTGAAG 13339
QY 9057 TCGAGCTACCGGATGCCCGCAGCTTCTGCGGAAAGATGTCCTCCCTCCCTAGTTT 9116
DB 13340 TCGAGCTACCGGATGCCCGCAGCTTCTGCGGAAAGATGTCCTCCCTCCCTAGTTT 13399
QY 9117 CTGTAAACCCAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCGCATCTCGAA 9176
DB 13400 CTGTAAACCCAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCGCATCTCGAA 13459
QY 9177 GGGGAAGCAGAGCCCTCTCGGCGCAGCAAGGCCCTCTGACTTCAGGGATAGACTCTAG 9236
DB 13460 GGGGAAGCAGAGCCCTCTCGGCGCAGCAAGGCCCTCTGACTTCAGGGATAGACTCTAG 13519
QY 9237 CACTTTAGCCCCCAAGTAACTGAGTCCACAGACACACAGCAATGTGGGGTCCAAG 9296
DB 13520 CACTTTAGCCCCCAAGTAACTGAGTCCACAGACACACAGCAATGTGGGGTCCAAG 13579
QY 9297 GCGCGCACCCAGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCCCTGCACC 9356
DB 13580 GCGCGCACCCAGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCCCTGCACC 13639
QY 9357 GGATAAGCAAAAAGACCCGAGACAAAATTTGAAACCTTTCAGTAGTCTTCTCAGCC 9416
DB 13640 GGATAAGCAAAAAGACCCGAGACAAAATTTGAAACCTTTCAGTAGTCTTCTCAGCC 13699
QY 9417 CACCTTCGCGCAGTAAATCATCATGTTTCCATCGACTACCCGCCCAAGACAGCC 9476
DB 13700 CACCTTCGCGCAGTAAATCATCATGTTTCCATCGACTACCCGCCCAAGACAGCC 13759
QY 9477 GTGGGCACTTCGCGGAGAGACTATTTCTGTATCGAGTCCGAGCAATACTTCAGCAGA 9536
DB 13760 GTGGGCACTTCGCGGAGAGACTATTTCTGTATCGAGTCCGAGCAATACTTCAGCAGA 13819
QY 9537 ATTTGTGAGCCCTTCCAAAAGACTGAAGTCCAGCGTGTGTCCTTGAAGTTGACCA 9596
DB 13820 ATTTGTGAGCCCTTCCAAAAGACTGAAGTCCAGCGTGTGTCCTTGAAGTTGACCA 13879
QY 9597 GCGCGGGCCNATTACAGAGAGGCTATGACCTTCCAGTACCATATGTCAGGCGAT 9656
DB 13880 GCGCGGGCCNATTACAGAGAGGCTATGACCTTCCAGTACCATATGTCAGGCGAT 13939
QY 9657 CACATCACTGTTTACCAGGACTGTGTATCGTCCGAGGCGCTCCCTCCCAACCAAG 9716
DB 13940 CACATCACTGTTTACCAGGACTGTGTATCGTCCGAGGCGCTCCCTCCCAACCAAG 13999
QY 9717 GTTCTCTGAGCTCAGCGAGTGGATTTCTCAAAATGTCGTGACTGTTTCAAGGCCCTTATCG 9776
DB 14000 GTTCTCTGAGCTCAGCGAGTGGATTTCTCAAAATGTCGTGACTGTTTCAAGGCCCTTATCG 14059
QY 9777 TGCTCCGGGCCACTTTTACATGTTGTGCTGTGTGTATCGTCCGAGGCGCTCCCTCCCAACCAAG 9836
DB 14060 TGCTCCGGGCCACTTTTACATGTTGTGCTGTGTGTATCGTCCGAGGCGCTCCCTCCCAACCAAG 14119
QY 9837 AGAAGGATTTGCAATGAGGCGCTGCTGTTTAAATGCTGCTACAGTGAATTAATGCTAA 9896
DB 14120 AGAAGGATTTGCAATGAGGCGCTGCTGTTTAAATGCTGCTACAGTGAATTAATGCTAA 14179
QY 9897 TCCAGGCAATTTCTAGTGGAGATGGTACCACCTCCCAAGGGTGGGGGTAGGCGCCAGAA 9956

DB 14180 TCCAGGCAATTTCTAGTGGAGATGGTACCACCTCCCAAGGGTGGGGGTAGGCGCCAGAA 14239
QY 9957 TTCTTGGGGTTCACAGAGAAAGCATTTCTAGTATCGGCAAGTGG 10000
DB 14240 TTCTTGGGGTTCACAGAGAAAGCATTTCTAGTATCGGCAAGTGG 14283

RESULT 5
AC015713/C
LOCUS
DEFINITION
Homo sapiens chromosome 22 clone RP11-67H19 map 22, WORKING DRAFT
SEQUENCE, 24 unordered pieces.
AC015713
AC015713.3 GI:7329268
HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 165471)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 22, clone RP11-67H19
Unpublished
2 (bases 1 to 165471)
REFERENCE
1 (bases 1 to 165471)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Boukhgalter, B.,
Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,
Cooke, P., DeArelano, K., Dewar, K., Domino, M., Donegan, L., Doyle, M.,
Ferrel, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,
Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lehoczky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J.,
Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
3 (bases 1 to 165471)
Autors
Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 165471)
Autors
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 26, 2000 this sequence version replaced gi:6514042.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

TITLE
JOURNAL
COMMENT
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 26, 2000 this sequence version replaced gi:6514042.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information -----
Center project name: L1398
Center clone name: 67.H.19
----- Summary Statistics -----
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 155715 bases at least Q40
Consensus quality: 160515 bases at least Q30
Consensus quality: 162183 bases at least Q20
Insert size: 138000; agarose- ϕ p
Insert size: 163171; sum-of- ϕ -contigs
Quality coverage: 6.8 in Q20 bases; agarose- ϕ p
Quality coverage: 5.8 in Q20 bases; sum-of- ϕ -contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1223: contig of 1323 bp in length
* 1324 1423: gap of 100 bp
* 1424 2727: contig of 1304 bp in length
* 2728 2827: gap of 100 bp
* 2828 4344: contig of 1517 bp in length
* 4345 4444: gap of 100 bp
* 4445 5722: contig of 1278 bp in length
* 5723 5822: gap of 100 bp
* 5823 6884: contig of 1062 bp in length
* 6885 6984: gap of 100 bp
* 6985 8290: contig of 1306 bp in length
* 8291 8390: gap of 100 bp
* 8392 9932: contig of 1542 bp in length
* 9933 10032: gap of 100 bp
* 10033 12359: contig of 2327 bp in length
* 12360 12459: gap of 100 bp
* 12460 14959: contig of 2500 bp in length
* 14960 15059: gap of 100 bp
* 15060 18997: contig of 3938 bp in length
* 18998 19097: gap of 100 bp
* 19098 25489: contig of 6392 bp in length
* 25490 25589: gap of 100 bp
* 25590 31287: contig of 5698 bp in length
* 31288 31387: gap of 100 bp
* 31388 36301: contig of 4914 bp in length
* 36302 36401: gap of 100 bp
* 36402 42560: contig of 6159 bp in length
* 42561 42660: gap of 100 bp
* 42661 48389: contig of 5729 bp in length
* 48390 48489: gap of 100 bp
* 48490 54499: contig of 6009 bp in length
* 54499 54598: gap of 100 bp
* 54599 61031: contig of 6433 bp in length
* 61032 61131: gap of 100 bp
* 61132 68382: contig of 7251 bp in length
* 68383 68482: gap of 100 bp
* 68483 78452: contig of 9870 bp in length
* 78453 78552: gap of 100 bp
* 78453 89588: contig of 11136 bp in length
* 89589 89688: gap of 100 bp
* 89689 100432: contig of 10744 bp in length
* 100433 100532: gap of 100 bp
* 100533 112962: contig of 12430 bp in length
* 112963 113062: gap of 100 bp
* 113063 134930: contig of 21868 bp in length
* 134931 135030: gap of 100 bp
* 135031 165471: contig of 30441 bp in length.

FEATURES
source

Location/Qualifiers
1..165471
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="22"
/map="22"
/clone="RP11-67H19"
/clone_lib="RPC1-11 Human Male BAC"
1..1323
/note="assembly_fragment"
1424..2727
/note="assembly_fragment"
2828..4344
/note="assembly_fragment"
4445..5722
/note="assembly_fragment"
5823..6884
/note="assembly_fragment"
6985..8290
/note="assembly_fragment"
8391..9932
/note="assembly_fragment"
10033..12359
/note="assembly_fragment"
12460..14959
/note="assembly_fragment"
clone_end:SP6
vector_side:right
15060..18997
/note="assembly_fragment"
19098..25489
/note="assembly_fragment"
25590..31287
/note="assembly_fragment"
31388..36301
/note="assembly_fragment"
clone_end:T7
vector_side:right
36402..42560
/note="assembly_fragment"
42661..48389
/note="assembly_fragment"
48490..54498
/note="assembly_fragment"
54599..61031
/note="assembly_fragment"
61132..68382
/note="assembly_fragment"
68483..78352
/note="assembly_fragment"
78453..89588
/note="assembly_fragment"
89689..100432
/note="assembly_fragment"
100533..112962
/note="assembly_fragment"
113063..134930
/note="assembly_fragment"
135031..165471
/note="assembly_fragment"

ORIGIN

Query Match 21.3%; Score 2132; DB 2; Length 165471;
Best Local Similarity 96.6%; Pred.No. 0;
Matches 2248; Conservative 0; Mismatches 55; Indels 23; Gaps 6;

QY 1 CCATCATATTTCTTTATTTTGGCGGAGAGGGG-----AGATTGCTCTGTGCCCA 54
|||||
DB 67295 CCATCATATTTCTTTTGGCGAAGAGGGGGAGTCAAAGTCTGCTGTGCCCA 67166
|||||
QY 55 GGCTGGA--CCAGTGGTGGATCTTGGGTCACTGCAACCTCCACCTCCTGGGTCAAGTG 112
|||||

repeat_region	/rpt_family="L1"	46673..46956	QY	154	CAGCTAATTTTGTATTTTGTAGCAGATAAGGGGTTTCCACCATGTTGSCCAGGCTGGTCTC	213
repeat_region	/rpt_family="ALU"	complement(47515..48025)	Db	37549	CAGCTAATTTTGTATTTTGTAGCAGATAAGGGGTTTCCACCATGTTGSCCAGGCTGGTCTC	37608
repeat_region	/rpt_family="MER42"	49234..49276	QY	214	CAACTCCTGGCCTCATGTGATCCACCACCTTCGGCTTCCAAAGCATTTGGAGATATAGGT	273
repeat_region	/rpt_family="THER1"	complement(52666..52717)	Db	37609	CAACTCCTGGCCTCAAGTGTATCCACCACCTTCGGCTACCCAAAGCATTTGGAGATATAGGT	37668
repeat_region	/rpt_family="MLT1"	53759..54035	QY	274	GTGAGCCACTATACCCGCTCTCATCATATATTTCTAATCCCGAGACTGTAGAGCTGGTGT	333
repeat_region	/rpt_family="ALU"	complement(55174..55236)	Db	37669	GTGAGCCACTATACCCGAGCTCTCATCATATATTTCTAATCCCGAGACTGTAGAGCTGGTGT	37728
repeat_region	/rpt_family="ALU"	complement(55782..56814)	QY	334	CTCTTTTCTTAAAGGATGTCAAGTAGAGAGTGGAGTTCCCAAAATTTACAGTTTTCAGTA	393
repeat_region	/rpt_family="L1"	59423..59715	Db	37729	CTCTTTTGGAAAGGATGTCAAGTAGAGAGTGGAGTTCCCAAAATTTACAGTTTTCAGTA	37788
repeat_region	/rpt_family="ALU"	61070..61335	QY	394	TTAGTCAAGTTTCTAAAAATACAGTAATAATGTTGAGAGCTGACATAGGGACTAACTTGGT	453
repeat_region	/rpt_family="ALU"	complement(61766..62009)	Db	37789	TTAGTAAAGTTTCTAAAAATACAGTAATAATGTTGAGAGCTGACATAGGGACTAACTTGGT	37847
repeat_region	/rpt_family="MER43"	complement(63650..67174)	QY	454	TTTTTTTTTTTTTTTTTTTTTTTTTTTCAAAATTTCTCACTGAACCTTTGATTTTGTCTAAATAGGAC	513
repeat_region	/rpt_family="L1"	67188..67468	Db	37848	TTTTTTTTTTTTTTTTTTTTTTTTTTTCAAAATTTCTCACTGAACCTTTGATTTTGTCTAAATAGGAC	37901
repeat_region	/rpt_family="ALU"	complement(70783..71062)	QY	514	ATTAAAAAACCACCAAACTCCACTATTGCTTATTCCTACTATTGCTTATTTGATTTTAAAAA	573
repeat_region	/rpt_family="ALU"	72529..72783	Db	37902	ATT-AAAAAACCACCAAAAGTCCACTATTGCTTATTCCTACTATTGCTTATTTGATTTTAAAAA	37960
repeat_region	/rpt_family="MER7"	complement(74296..74439)	QY	574	AATAAGCGTATTTTAGCATCTAAAAAGTAGGAAGGACCTCAAAATAAATAGAGTCTTTGTCT	633
repeat_region	/rpt_family="MER5"	75366..75612	Db	37961	AATAAGCGTATTTTAGCATCTAAAAAGTAGGAAGGACCTCAAAATAAATAGAGTCTTTGTCT	38020
repeat_region	/rpt_family="ALU"	75666..75943	QY	634	TGGCCAGGAAAAACAGCGTTGTGAGAAATTTGATACTGTTTCTTAGGGTATGTGCTGTT	693
repeat_region	/rpt_family="ALU"	78172..78402	Db	38021	TGGCCAGGAAAAACAGCGTTGTGAGAAATTTGATACTGTTTCTTAGGGTATGTGCTGTT	38080
repeat_region	/rpt_family="ALU"	complement(78597..79032)	QY	694	ATTCAAGTTAAAACTTGGCTGGGACGTAGCAATTCAGTAAATACTGTTGTTGAATAAGCAAA	753
repeat_region	/rpt_family="L1"	79128..79408	Db	38081	ATTCAAGTTAAAACTTGGCTGGGACGTAGCAATTCAGTAAATACTGTTGTTGAATAAGCAAA	38140
repeat_region	/rpt_family="ALU"	79926..80200	QY	754	TGAAACTTAAAGCTTCTATGTATAGAAACCTAAAGTCACTTTACATTTCTGATTTAGCAGATG	813
repeat_region	/rpt_family="ALU"	82675..89455	Db	38141	TGAAACTTAAAGCTTCTATGTATAGAAACCTAAAGTCACTTTACATTTCTGATTTAGCAGATG	38200
repeat_region	/rpt_family="L1"	92817..93048	QY	814	ATTGAATATTCTTTCAATGTGTAGCTCTATCCCGAGAACCAAGAGATATTTGGAACCTGTA	873
repeat_region	/rpt_family="ALU"	complement(93239..93306)	Db	38201	ATTGAATATTCTTTCAATGTGTAGCTCTATCCCGAGAACCAAGAGATATTTGGAACCTGTA	38260
repeat_region	/rpt_family="MER46"	complement(94031..94324)	QY	874	AAGGCCATCTCTATAGTTTAAACCAACTGCGTTAAATAGATAATAGAAAGATGTTGATGTC	933
repeat_region	/rpt_family="ALU"	complement(94710..94999)	Db	38261	AA-----AACCAACTGCAATTAAT-----AGATGTTGTTATGTG 38293	
repeat_region	/rpt_family="ALU"	complement(99945..100244)	QY	934	GCAGTGACAACTTGAAGGTTGTGACTAGAACTCGGGTCTCTCGAGTGTCTTATTTATATCA	993
repeat_region	/rpt_family="ALU"	100754..101034	Db	38294	GCAGTGACAACTTGAAGGTTGTGACTAGAACTCGGGTCTCTCGAGTGTCTTATTTATCA	38353
repeat_region	/rpt_family="ALU"	complement(104640..104845)	QY	994	CACCAAGCTGTGCACAGCCCATGTTGATCCTCCATCTGATAGCAACCAAGAGATGTTGATG	1053
repeat_region	/rpt_family="ALU"	complement(106967..111123)	Db	38354	CACCAAGCTGTGCACAGCCCATGTTGATCCTCCATCTGATAGCAACCAAGAGATGTTGATG	38413
gene	/gene="P8VFPp"		QY	1054	CTTCAGGACATTTCTTTTACCCCTTAATCTTGTATCTGCAGCTCTTATTTAGAAAGCTT	1113
			Db	38414	CTTCAGGACATTTCTTTTACCCCTTAATCTTGTATCTTTCAGTCTTATTTAGAAAGCTT	38473
			QY	1114	AATGTTAAAGATCTAGTTTATTTCAAAAATTAAGATAAAGAGATATGAGATTTCTATTT	1173
			Db	38474	AATGTTAAAGATCTAGTTTATTTCAAAAATTAAGAT---AAGAGATATGAGATTTCTATTT	38530
			QY	1174	TCGGAGTGTAAAGAGGAGATGTTTCTTTGGCTTCTCTGAGCCTGAGGCCCTTCTTGTCT	1233
			Db	38531	TCGGAGTGTAAAGAGGAGATGTTTCTTTGGCTTCTCTGAGCCTGAGGCCCTTCTTGTCT	38590

Query Match 20.2%; Score 2015; DB 9; Length 111123;
Best Local Similarity 95.3%; Pred.No. 0;
Matches 2178; Conservative 0; Mismatches 60; Indels 47; Gaps 8;
QY 36 AGACTTGTCTGTGTCAGGCTGGA--CCAGTGTGCGATCTTGGCTCACTGCAACCTC 93
Db 37429 AGTCTGCCCTGTGTCAGGCTGGAATGAGTGGCAGCATCTTGGCTCACTGCTACCTC 37488
QY 94 CACCTCTGGTTCAGATGATTTCCAAATAGCTGGGATACAGGTGTGATTACCATGCC 153
Db 37489 CACTTCTGGTTCAGTGTGATTTCCAAAGTAGTGGGATTTACAGGTGTGATTACCATGCC 37548

Matches 2178; Conservative 0; Mismatches 60; Indels 47; Gaps 8;				
QY	36	AGACTTCTCTGTTGCCAGGCTGGA--CCAGTGGTGGATCTTTGGCTCACTGCAACCTC	93	
Db	174350	AGTCTGCCCTGTTGGCCAGGCTGGAATGAGTGGCAGCATCTTTGGCTCACTGCTACCTC	174291	
QY	94	CACCTCTCGGGTTCAGTGAATCCAAATAGCTGGGATACAGGTGATATACCAATGCC	153	
Db	174290	CACCTCTCGGGTTCAGTGAATCCAAATAGCTGGGATACAGGTGATATACCAATGCC	174231	
QY	154	CAGCTAATTTTGTATTTTAGCAGATAAGGGTTTCCACCATGTTGGCCAGGCTGGTCTC	213	
Db	174230	CAGCTAATTTTGTATTTTAGCAGATAAGGGTTTCCACCATGTTGGCCAGGCTGGTCTC	174171	
QY	214	CAACTCTGGCCTCATGTGATCCACCACTTCGGCTTTCCAAAGCATTTGGGAGTATAGGT	273	
Db	174170	CAACTCTGGCCTCAAGTGATCCACCACCTTCGGCTACCCAAAGCATTTGGGAGTACAGGG	174111	
QY	274	GTGAGCCACTATACCGTCTCTCATCATATATTTCTTAATCCCGAGACTGTAGAGCTGGTGT	333	
Db	174110	GTGAGCCACTATGCCAGCCCTCATCATATATTTCTTAATCCCGAGACTGTAGAGCTGGTGT	174051	
QY	334	CTCTTTTCTAAAGGATGTCAGTAGAGAGTGGAGTTTCCCAAAATTTACAGTTTCAAGTA	393	
Db	174050	CTCTTTTGGAAAGGATGTCAGTAGAGAGTGGAGTTTCCCAAAATTTACAGTTTCAAGTA	173991	
QY	394	TTAGTCAAGTTTCTAAATAACAGTAATAATGTTGAGAGCTGACATAGGGACTAACTTGGT	453	
Db	173990	TTAGTAAAGTTTCTAAATAACAGTAATAATGTTGAGAGCTGACATAGGGACTAACTTAG-	173932	
QY	454	TTTTTTTTTTTTTTTTTTTTTCCAAATTTGTCAGTAACTTTGCTTAAATAAGGAC	513	
Db	173931	-----TTTTTTTTTTTTTTTTTCCAAATTTGTCAGTAACTTTGCTTAAATAAGGAC	173878	
QY	514	ATTAATAAAACCAACCAAAACCTCCACTATGCTTATGCTTATGCTTTTAAAA	573	
Db	173877	ATT-AAAGAAACCAACCAAAAGTCCACTATGCTTATGCTTATGCTTTTAAAA	173819	
QY	574	AATAAGCGTATTTAGCATCTAAAGTAGGAAGGACCTCAATAAATAGTCTTTGTTCT	633	
Db	173818	AATAAGCGTATTTAGCATCTAAAGTAGGAAGGACCTCAATAAATAGTCTTTGTTCT	173759	
QY	634	TGSCCAGGGAACACAGCTGTGTCAGAAATTTGATAAATGTTTCTAGGGTATGCTGTT	693	
Db	173758	TGSCCAGGGAACACAGCTGTGTCAGCAATTTGATAAATGTTTCTAGGGTATGCTGTT	173699	
QY	694	ATTGAGTTAAACCTTGGCCAGGCTAGCATTCAGTAAATACCTTTGTAATAACCAA	753	
Db	173698	ATTGAGTTAAACCTTGGCCAGGCTAGCATTCGGTAAATACCTTTGTAATAACCAA	173639	
QY	754	TGAACCTTAAGCTTCTATGTATAGAAACCTTAAGTCACTTCAATTTCTGATTAGCAGATTA	813	
Db	173638	TGAACCTTAAGCTTCTATGTATAGAAACCTTAAGTCACTTCAATTTCTGATTAGCAGATTA	173579	
QY	814	ATTGAATATTTCTTTCAATGTGATGCTCTATCCCCAGAACCAAGAAATATGGAACCTGTA	873	
Db	173578	ATTGAATATTTCTTTCAATGTGATGCTCTATCCCCAGAACCAAGAAATATGGAACCTGTA	173519	
QY	874	AAGGCCATCTATGTTTAAACCACTGCGTTAAATAGATAATAGAAAGATGTTGATGTG	933	
Db	173518	AA-----AACCACTGCAATTAAT-----AGATGTGGTATGTG	173486	
QY	934	GCAGTGACAACTTGAAGTGTGTGACTAGAACTCGGGTCTCTGGAGTGTCTATTATATCA	993	
Db	173485	GCAGTGACAACTTGAAGTGTGTGACTAGAACTCGGGTCTCTGGAGTGTCTATTATATCA	173426	
QY	994	CACCAAGCTGGTCAACAGCCCATGTGTTGATCTCTCATTTGTGATAGCAACAAAGAAAGA	1053	
Db	173425	CACCAAGCTGGTCAACAGCCCATGTGTTGATCTCTCATTTGTGATAGCAACAAAGAAAGA	173366	
QY	1054	CTTCAGGACATTTCTTTCCCTTTTACCCTTAATCCTTTGATCTGCAGTCTTATTTAGAAAGCTT	1113	
Db	173365	CTTCAGGACATTTCTTTCCCTTTTACCCTTAATCCTTTGATCTTTCACTTTATTTAGAAAGCTT	173306	

QY	1114	AATGTTTAAAGATCTAGTATTTTAACTAAAGATAACAAGGAGTATGAGAAATTTCTATT	1173	
Db	173305	AATGTTTAAAGATCTAGTATTTTAACTAAAGAT---AAGGAGTATGAGAAATTTCTATT	173249	
QY	1174	TCGGAGTGTAAAGAGGAGATGTTTCTTGCTTCTCTGAGCCTGACAGGCTTCTCTTGTCT	1233	
Db	173248	TCGGAGTGTAAAGAGGAGATGTTTCTTGCTTCTCTGAGCCTTACAGGCTTCTCTTGTCT	173189	
QY	1234	CTTTAAGGAGTATAGAGAGGAGGAGAAAGTAAAGTATGCTTTTGTTTTTTAAGGTACTT	1293	
Db	173188	CTTTAAGGAGTATAGAGAGGAGGAGAAAGTAAAGTATGCTTTTGTTTTTTAAGGTACTT	173129	
QY	1294	TGCTGGGAGTATGTCATGCCCTTTTGGTCTTCTTGCTGCGAATTAACCTTAACTT	1353	
Db	173128	TGCTGGGAGTATGTCATGCCAATTTGGTCTTCTTGCTGCGAATTAACCTTAACTT	173069	
QY	1354	TAAATAGTTGGGACTATTTTAAAAAACAATGCTATCCAATGTTTGGCCATAAAGCAGAGG	1413	
Db	173068	TAAATAGTTGGGACTATTTTAAAAAACAATGCTATCCAATGTTTGGCCATAAAGCAGAGG	173009	
QY	1414	TATTGGCTTTAGAGGTTAAATCTTCTCCAGGAGTGAATTTAGCTTCTAAACCAGAGCA	1473	
Db	173008	TATTGGCTTTAGAGGTTAAATCTTCTCCAGGAGTGAATTTAGCTTCTAAACCAGAGCA	172949	
QY	1474	GCAGACTAAATAAAGTAAATTTTCCACCTGGCCAGTGCATGATGTAAGAGGTAGATTTAA	1533	
Db	172948	GCAGACTAAAGAAAGTAAATTTTCCACCTGGCCAGTGCATGATGTAAGAGGTAGATTTAA	172889	
QY	1534	AAAAAGAGAGGCCCCATTTTCTGATGAAAGACATAAGCCATGTTTGAACACGCCCTGTGAG	1593	
Db	172888	AGAAATGAGAGGCCCCATTTTCTGATGAAAGACATAAGCCATGTTTGAACACGCCCTGTGAG	172829	
QY	1594	GATTTTATTTTAAATCTATACATTTCAAAAGGAGCTTTGTGATGCTTTTCCCTATTGT	1653	
Db	172828	GATTTTATTTTAAATCTATACATTTCAAAAGGAGCTTTGTGATGCTTTTCCCTATTGT	172769	
QY	1654	TGTTTGGACTAGGAAGCCCCACCCAGTGTCTGTTGAAGCAGAAAGTCGTTGAAAGCAAG	1713	
Db	172768	TGTTTGGACTAGGAAGCCCCACCCAGTGTCTGTTGAAGCAGAAAGTCGTTGAAAGCAAT	172709	
QY	1714	CTGGGATTTGAAACAGTGGATTCAGGTTTCCAAATATCCAGTGAACCAAAATATATCAGGGT	1773	
Db	172708	CCGGATTTTGAACAGTGGATTCAGGTTTCCAAATATCCAGTGAACCAAAATATATCAGGGT	172649	
QY	1774	TCCCTTGGCCAGATGAGTGACCATCTCGAGTGTTAAGTATTTCTTGAATGGGATTTT	1833	
Db	172648	TCCCTTGGCCAGATGAGTGACCATCTCGAGTGTTAAGTATTTCTTGAATGGGATTTT	172589	
QY	1834	AGGAAAAGTTTCTGTATTTCTGTGCTCATTTTGTGACCTCTGTATGTGCAAAATCTCTA	1893	
Db	172588	AGGAAAAGTTTCTGTATTTCTGTGCTCATTTTGTGACCTCTGTATGTGCAAAATCTCTA	172529	
QY	1894	AGGGGTGTTTGGGCACTTAGATTTCTTGGATTCAGATTTGTTTGTATATGAACAAAT	1953	
Db	172528	AGGGGTGTTTGGGCACTTAGATTTCTTGGATTCAGATTTGTTTGTATATGAACAAAT	172469	
QY	1954	TTAAATTTGTTTGTATACACTCGATTTTAAA-----TAGTTTACTTAAAGTGTTTTAA	2008	
Db	172468	TTAAATTTGTTTGTATACACTCGATTTTAAAATAGTTTACTTAAAGTGTTTTAA	172409	
QY	2009	TTTCACTTAAATTTTTCAGAGTTCTTATAGTCTTTTAAATTTAGGAGGCTGTTGATGGCAT	2068	
Db	172408	TTTCACTTAAATTTTTCAGTTCTTATAGTCTTTTAAATTTAGGAGGCTGTTGATGGCAT	172351	
QY	2069	CCACATGTGCAATTTTGTGGCAATTTTAAATATGATTTTCACTGAAATTTTAACTTCTGACC	2128	
Db	172350	CCACATGTGCAATTTTGTGGCAATTTTAAATATGATTTTCACTGAAATTTTAACTTCTGACC	172291	
QY	2129	TAAACTTTGACATTTTGTAGATTTTAAAGTCGTAAGCACTGATTTTAACTGGATTTTAACTG	2188	
Db	172290	TAAACTTTGACATTTTGTAGATTTTAAAGTCGTAAGCACTGATTTTAACTGGATTTTAACTG	172231	

```
QY 2189 GATGAATTTCTGATTTAATAAGTACTGACTGGATAAATGCCAATGATTTAATTAAACA 2248
DB 172230 GATGAATTTCTGATTTAATAAGTACTGACTGGATAAATGCCAATGATTTAATTAAACA 172171
QY 2249 AGCAGCTTTAAACAGGATGCCCTATATATATTAGTTAAAGTGAAGCAATTTGAATTAGGTACC 2308
DB 1721170 AGCAGCTTTAAACAGGATGCCCTATATATATTAGTTAAAGTGAAGCAATTTGAATTAGGTACC 172111
QY 2309 TTCTC 2313
DB 172110 ATTTCC 172106

RESULT 8
LOCUS AR157086 5632 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 3 from patent US 6242590.
ACCESSION AR157086
VERSION AR157086.1 GI:15125790
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 5632)
TITLE Cowser, L.M.
JOURNAL Antisense modulation of zinc finger protein-217 expression
FEATURES
    source 1..5632
    /organism="unknown"
    /mol_type="unassigned DNA"

ORIGIN
Query Match 16.4%; Score 1639.4; DB 6; Length 5632;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1640; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2474 GACAAAGAGAACTAATGCTTTTGCTGATTCATATTTGAATCGAGGCATTTGGGAACCCCTG 2533
DB 1 GACAAAGAGAACTAATGCTTTTGCTGATTCATATTTGAATCGAGGCATTTGGGAACCCCTG 60
QY 2534 TATGCTTTGTTGCGAAGAACCACTGACACCATCTGAGCTTCTTAAAGATTCGAAG 2593
DB 61 TATGCTTTGTTGCGAAGAACCACTGACACCATCTGAGCTTCTTAAAGATTCGAAG 120
QY 2594 AAGTTAGAGGACTATACACTTTCTTTTGAACCTTTTATAATAATATTTGCTCTGTTTTT 2653
DB 121 AAGTTAGAGGACTATACACTTTCTTTTGAACCTTTTATAATAATATTTGCTCTGTTTTT 180
QY 2654 GGAAACCCAGGCTGTTAGAGGGGTGAGTGACAAGCTTTACAAGTGGCCCTTATTCGAATC 2713
DB 181 GGAAACCCAGGCTGTTAGAGGGGTGAGTGACAAGCTTTACAAGTGGCCCTTATTCGAATC 240
QY 2714 CAGAAATGCCCAACGGAACTTTGAGATTTATGCAATCGAAGTGACAGGAAACATGCC 2773
DB 241 CAGAAATGCCCAACGGAACTTTGAGATTTATGCAATCGAAGTGACAGGAAACATGCC 300
QY 2774 AACTCAATCCCTTTAATGTACATGGATGGGCGCAGAGTGATTTGGCAGCTCTCTTTGGCAG 2833
DB 301 AACTCAATCCCTTTAATGTACATGGATGGGCGCAGAGTGATTTGGCAGCTCTCTTTGGCAG 360
QY 2834 TCCGATGGAGATGGAGGATGCTTGTCAATGAAGAGGACCGCTGTTTCCATTCGGAGC 2893
DB 361 TCCGATGGAGATGGAGGATGCTTGTCAATGAAGAGGACCGCTGTTTCCATTCGGAGC 420
QY 2894 TACACAAGAAAAATGTCTATCCAAATCGAGGGGTATATGCCCTTGGATTGATGTTCTG 2953
DB 421 TACACAAGAAAAATGTCTATCCAAATCGAGGGGTATATGCCCTTGGATTGATGTTCTG 480
QY 2954 CAGCCAGACCTTTCACACATTCGAGAAGACCTTAAATAACATGCTTTAATGCAACACCGGCC 3013
DB 481 CAGCCAGACCTTTCACACATTCGAGAAGACCTTAAATAACATGCTTTAATGCAACACCGGCC 540
```

```
QY 3014 TACCCTCTGTGAACACAGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 3073
DB 541 TACCCTCTGTGAACACAGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAA 600
QY 3074 AAGTCAAGTCGGAACAGAACCTTCCCAAGGAAAAAATTTGCAAGGAAAAATGAATTTAGCTG 3133
DB 601 AAGTCAAGTCGGAACAGAACCTTCCCAAGGAAAAAATTTGCAAGGAAAAATGAATTTAGCTG 660
QY 3134 TGAGGTATGTGGGCGAGACATTTAGAGTCGCTTTTGTATTTGAGATCCACATGAGAACACA 3193
DB 661 TGAGGTATGTGGGCGAGACATTTAGAGTCGCTTTTGTATTTGAGATCCACATGAGAACACA 720
QY 3194 CAAAGATTTCTTCACTTACGGGTGTAAATGTGCGGAAGAGATTTCAAGGAGCCTTGGTT 3253
DB 721 CAAAGATTTCTTCACTTACGGGTGTAAATGTGCGGAAGAGATTTCAAGGAGCCTTGGTT 780
QY 3254 TCTTAAAAATCACATGCGGACACATAATGCGAAATCGGGGGCCAGAAGCAAACTCGAGCA 3313
DB 781 TCTTAAAAATCACATGCGGACACATAATGCGAAATCGGGGGCCAGAAGCAAACTCGAGCA 840
QY 3314 AGGCTTGAGAGTAGTCCAGCAACATCAACGAGGTCTGTCAGGTGCAACGCGGCCGAGAG 3373
DB 841 AGGCTTGAGAGTAGTCCAGCAACATCAACGAGGTCTGTCAGGTGCAACGCGGCCGAGAG 900
QY 3374 CATCTCTCTCTTACAAAATCTGCATGGTTTGTGGCTTCTCTATTTCCAAAATAAAGAAAG 3433
DB 901 CATCTCTCTCTTACAAAATCTGCATGGTTTGTGGCTTCTCTATTTCCAAAATAAAGAAAG 960
QY 3434 TCTAATTTGAGCACCGCAAGGTGCACACCAAAAAAATCTGTTTCGGTACACAGCGCGCGCA 3493
DB 961 TCTAATTTGAGCACCGCAAGGTGCACACCAAAAAAATCTGTTTCGGTACACAGCGCGCGCA 1020
QY 3494 GACAGACTCTCCAAAGGAGGAATGCGCTCTCGAGGGAGGACTTCTTCGAGTTGTTTCA 3553
DB 1021 GACAGACTCTCCAAAGGAGGAATGCGCTCTCGAGGGAGGACTTCTTCGAGTTGTTTCA 1080
QY 3554 CTTGAGACCAAAATCTCACCTGAAAACGGGGGAGGAAGCTGTGCATGCATCCCTCAGCT 3613
DB 1081 CTTGAGACCAAAATCTCACCTGAAAACGGGGGAGGAAGCTGTGCATGCATCCCTCAGCT 1140
QY 3614 CGATCCGTTTCAACCTTCCAGGCTTGGCAGCTGGCTTACCAAGGAAAAATTTGCCATTTG 3673
DB 1141 CGATCCGTTTCAACCTTCCAGGCTTGGCAGCTGGCTTACCAAGGAAAAATTTGCCATTTG 1200
QY 3674 CCAAGAGTGAAGGAATCGGGGCAAGAGGAGGACCCGACACGACGATTCGAGTTCCGA 3733
DB 1201 CCAAGAGTGAAGGAATCGGGGCAAGAGGAGGACCCGACACGACGATTCGAGTTCCGA 1260
QY 3734 GAAAGAGCTTGGAGAAACAAATTAAGGCGAGTTGTGCGAGGCCCTCTCGCAAGAGAAAGAGAA 3793
DB 1261 GAAAGAGCTTGGAGAAACAAATTAAGGCGAGTTGTGCGAGGCCCTCTCGCAAGAGAAAGAGAA 1320
QY 3794 GTGCAAAACATCTCCACCGCGAAGCGCCCTCGGTGAGCGCGATCCCAAGTTACCCAGTAG 3853
DB 1321 GTGCAAAACATCTCCACCGCGAAGCGCCCTCGGTGAGCGCGATCCCAAGTTACCCAGTAG 1380
QY 3854 CAAAGAGAGGCCCACTCACTGCTCCGAGTGGGCAAGGTTTTCAGAACCTTACCAACAGCT 3913
DB 1381 CAAAGAGAGGCCCACTCACTGCTCCGAGTGGGCAAGGTTTTCAGAACCTTACCAACAGCT 1440
QY 3914 GGTCTTGACATCCAGGGTCCCAAGAGGACCGGAGGGCCGCGGAGTCCGCCACCAT 3973
DB 1441 GGTCTTGACATCCAGGGTCCCAAGAGGACCGGAGGGCCGCGGAGTCCGCCACCAT 1500
QY 3974 GTCTGTGAACGGGAGCGCGGAGCGTGTCTCTGACCTTCGCGGCCCTCTCGATGA 4033
DB 1501 GTCTGTGAACGGGAGCGCGGAGCGTGTCTCTGACCTTCGCGGCCCTCTCGATGA 1560
QY 4034 AAATGGAGCGGTGGATTCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 4093
DB 1561 AAATGGAGCGGTGGATTCGAGGGGAAGGTGGTTCTGAAGACGGATCTGAGGATGGGCTTCC 1620
QY 4094 CGAAGGAATCCATCTGGGTAA 4114
```



```
QY 3134 TGAGGTATGTGGGACAGATTTAGAGTCGCTTTTGTGATGTGATGCCATGAGAACACA 3193
Db 661 TGAGGTATGTGGGACAGATTTAGAGTCGCTTTTGTGATGTGATGCCATGAGAACACA 720
QY 3194 CAAGATTTCTTTCACCTTACCGGTGTAAATGTGCGGAAGAGATTTCAAGGACCTTGGTT 3253
Db 721 CAAGATTTCTTTCACCTTACCGGTGTAAATGTGCGGAAGAGATTTCAAGGACCTTGGTT 780
QY 3254 TCTTAAATACATGCGGACACATAATGGCAAAATCGGGGCCAGAACTCCAGCA 3313
Db 781 TCTTAAATACATGCGGACACATAATGGCAAAATCGGGGCCAGAACTCCAGCA 840
QY 3314 AGGCTTGGAGTAGTCCAGCAACGATCAACAGGTCTGTCAGGTGTCAGCGGCCGAGAG 3373
Db 841 AGGCTTGGAGTAGTCCAGCAACGATCAACAGGTCTGTCAGGTGTCAGCGGCCGAGAG 900
QY 3374 CATCTCCTCTCTTCAAAATCTGATGTTGTGCTTCTTCTTATTTCCAAATAAAGAAAG 3433
Db 901 CATCTCCTCTCTTCAAAATCTGATGTTGTGCTTCTTCTTATTTCCAAATAAAGAAAG 960
QY 3434 TCTAATTTAGACCCGCAAGGTGCACACCAAAAAAACTGCTTTTCTGTCAGCGCGCA 3493
Db 961 TCTAATTTAGACCCGCAAGGTGCACACCAAAAAAACTGCTTTTCTGTCAGCGCGCA 1020
QY 3494 GACAGACTCTCCACAGGAGGAAATGCGCTCTCGAGGGAGGACTTCTGCGAGTTGTTCAA 3553
Db 1021 GACAGACTCTCCACAGGAGGAAATGCGCTCTCGAGGGAGGACTTCTGCGAGTTGTTCAA 1080
QY 3554 CTTGAGACCAAAATCTCACCTTGAACCGGGGAAGAGCTGTGAGATGATCCCTCAGCT 3613
Db 1081 CTTGAGACCAAAATCTCACCTTGAACCGGGGAAGAGCTGTGAGATGATCCCTCAGCT 1140
QY 3614 CGATCCGTTTACCACCTTCCAGGCTTGGCAGCTGCTACCAAGGAAAGTTGCCATTG 3673
Db 1141 CGATCCGTTTACCACCTTCCAGGCTTGGCAGCTGCTACCAAGGAAAGTTGCCATTG 1200
QY 3674 CCAAGAGTGAAGGAATCGGGCAGAGAGGAGGACCGCAACAGCAGATTCAGTTCCGA 3733
Db 1201 CCAAGAGTGAAGGAATCGGGCAGAGAGGAGGACCGCAACAGCAGATTCAGTTCCGA 1260
QY 3734 GAAGGAGCTTGAGAGAAACAAATTAAGGCGAGTTGTGAGGCTCTCGAGGAGGAGAA 3793
Db 1261 GAAGGAGCTTGAGAGAAACAAATTAAGGCGAGTTGTGAGGCTCTCGAGGAGGAGAA 1320
QY 3794 GTGCAAAACATCCACCGGAGAGCGCTTCCGTGAGCGGATCCCAAGTTACCCAGTAG 3853
Db 1321 GTGCAAAACATCCACCGGAGAGCGCTTCCGTGAGCGGATCCCAAGTTACCCAGTAG 1380
QY 3854 CAAGGAGAGCCACTCACTGCTCCGAGTGGCGGCAAGCTTTTCAGAACCTTACCACAGCT 3913
Db 1381 CAAGGAGAGCCACTCACTGCTCCGAGTGGCGGCAAGCTTTTCAGAACCTTACCACAGCT 1440
QY 3914 GGTCTTGCACTCAGGCTCCACAGAGGACCGGAGGCGCGGAGTCCGCCACCAT 3973
Db 1441 GGTCTTGCACTCAGGCTCCACAGAGGACCGGAGGCGCGGAGTCCGCCACCAT 1500
QY 3974 GTCTGTGAGCGGAGGACCGGAGCGTGTCTCTGACCTTCGCGGCCCTCTCGATGA 4033
Db 1501 GTCTGTGAGCGGAGGACCGGAGCGTGTCTCTGACCTTCGCGGCCCTCTCGATGA 1560
QY 4034 AAATCGAGCGTGGATCGAGGGAGAGTGTCTGAGACGATCTGAGGATGGGCTCC 4093
Db 1561 AAATCGAGCGTGGATCGAGGGAGAGTGTCTGAGACGATCTGAGGATGGGCTCC 1620
QY 4094 CGAAGGAATCCATCTGGGTAA 4114
Db 1621 CGAAGGAATCCATCTGGATA 1641
```

RESULT 10
CQ726295
LOCUS

CQ726295 5633 bp DNA linear PAT 03-FEB-2004

```
DEFINITION Sequence 12229 from Patent WO02068579.
ACCESSION CQ726295
VERSION CQ726295.1 GI:42288937
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 12229 06-SEP-2002;
PE Corporation (US)
FEATURES
source
Location/Qualifiers
1..5633
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 16.4%; Score 1639.4; DB 6; Length 5633;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1640; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2474 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCAATGGGAACCTG 2533
Db 1 GACAAAGAGAACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCAATGGGAACCTG 60
QY 2534 TATGCTGTTTGTGGAAGAACCACTGACACCATCTGAGTTCCTTAAAGATTCGAAG 2593
Db 61 TATGCTGTTTGTGGAAGAACCACTGACACCATCTGAGTTCCTTAAAGATTCGAAG 120
QY 2594 AAGTTAGAGGACTATACACTTTCTTTTGAACCTTTTATAATAATTTGCTCTGTTTTT 2653
Db 121 AAGTTAGAGGACTATACACTTTCTTTTGAACCTTTTATAATAATTTGCTCTGTTTTT 180
QY 2654 GGAACCCAGGGCTGTAGAGGGTGAGTGACAAAGTCTTACAAAGTGGCTTTATTCACATC 2713
Db 181 GGAACCCAGGGCTGTAGAGGGTGAGTGACAAAGTCTTACAAAGTGGCTTTATTCACATC 240
QY 2714 CAGAAATTTGCCAACCGAACTTTGAGATTTATGCAATCGAAAGTGACAGGAAACATGCC 2773
Db 241 CAGAAATTTGCCAACCGAACTTTGAGATTTATGCAATCGAAAGTGACAGGAAACATGCC 300
QY 2774 AACTCAATCCCTCTTAATGTACATGGATGGGCGCAGAGTGATTTGGCAGCTCTTTGGCAG 2833
Db 301 AACTCAATCCCTCTTAATGTACATGGATGGGCGCAGAGTGATTTGGCAGCTCTTTGGCAG 360
QY 2834 TCCGATGAGATGGAGATGCTTGTCAATGAAAGGACCGCTGTGTTTCCATTCGAGC 2893
Db 361 TCCGATGAGATGGAGATGCTTGTCAATGAAAGGACCGCTGTGTTTCCATTCGAGC 420
QY 2894 TACACAGCAAAAAATGTCAATCGAGGGTATATGCCCTTGGATTGATGTTCTG 2953
Db 421 TACACAGCAAAAAATGTCAATCGAGGGTATATGCCCTTGGATTGATGTTCTG 480
QY 2954 CAGCGAGACCTTTCACACATTCAGAGAGCTTTAATAAACAATGCTTTAATGCAACACCGGCC 3013
Db 481 CAGCGAGACCTTTCACACATTCAGAGAGCTTTAATAAACAATGCTTTAATGCAACACCGGCC 540
QY 3014 TACCTCTGTGAAACAGCAGTTTTCGGGTGGAACAGAGTATCTCAGTCCGCTTGATTA 3073
Db 541 TACCTCTGTGAAACAGCAGTTTTCGGGTGGAACAGAGTATCTCAGTCCGCTTGATTA 600
QY 3074 AAGTCAAGTGGGAACAGAACCTTCCCAAGGAAAGAAATTCGAAGGAAATGAATTTAGCTG 3133
Db 601 AAGTCAAGTGGGAACAGAACCTTCCCAAGGAAAGAAATTCGAAGGAAATGAATTTAGCTG 660
QY 3134 TGAGGTATGTGGGACAGATTTAGAGTCGCTTTTGTGATGTGAGATCCACATGAGAACACA 3193
Db 661 TGAGGTATGTGGGACAGATTTAGAGTCGCTTTTGTGATGTGAGATCCACATGAGAACACA 720
```

```
QY 3194 CAAAGATTCTTTCACTTACCGGTGTAAATGTCGGAAGAAGATTCAAGAGAGCCTTGGTT 3253
Db 721 CAAAGATTCTTTCACTTACCGGTGTAAATGTCGGAAGAAGATTCAAGAGAGCCTTGGTT 780
QY 3254 TCTTAAATACATGCGGACACATATGCAATCGGGGCCGAAGCAAACTGCGACGA 3313
Db 781 TCTTAAATACATGCGGACACATATGCAATCGGGGCCGAAGCAAACTGCGACGA 840
QY 3314 AGGCTTGGAGTAGTACGACACGATCAACAGAGTCTGTCAGGTGACGCGCGCGAGAG 3373
Db 841 AGGCTTGGAGTAGTACGACACGATCAACAGAGTCTGTCAGGTGACGCGCGCGAGAG 900
QY 3374 CATCTCTCTCTTACAAAATCTGCATGTTGTGGTCTCTATTTCCAAATAAAGAAAG 3433
Db 901 CATCTCTCTCTTACAAAATCTGCATGTTGTGGTCTCTATTTCCAAATAAAGAAAG 960
QY 3434 TCTAAATTGAGCAGCGCAAGGTGACACCAAAAATCTGCTTTCGTACGAGCGCGCA 3493
Db 961 TCTAAATTGAGCAGCGCAAGGTGACACCAAAAATCTGCTTTCGTACGAGCGCGCA 1020
QY 3494 GACAGACTCTCCACAGGAGGAATGCGCTCTCGAGGAGACTTCTGCGAGTTGTTCAA 3553
Db 1021 GACAGACTCTCCACAGGAGGAATGCGCTCTCGAGGAGACTTCTGCGAGTTGTTCAA 1080
QY 3554 CTTGAGACCAAAATCTCACCTGAAACGGGGAAGAAGCCTGTGATGCAATCCCTCAGCT 3613
Db 1081 CTTGAGACCAAAATCTCACCTGAAACGGGGAAGAAGCCTGTGATGCAATCCCTCAGCT 1140
QY 3614 CGATCGTTTACACACTTCCAGGCTTGGAGCTTGGAGCTGCTCAAAAGGAAGAGTTGCCATTTG 3673
Db 1141 CGATCGTTTACACACTTCCAGGCTTGGAGCTTGGAGCTGCTCAAAAGGAAGAGTTGCCATTTG 1200
QY 3674 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCACAAAGAGATTCGAGTTCCGA 3733
Db 1201 CCAAGAAGTGAAGGAATCGGGGCAAGAAGGGAGCACCACAAAGAGATTCGAGTTCCGA 1260
QY 3734 GAAGGAGCTTGGAGAAACAAATTAAGGCGAGTTGTGAGGCGCTCTCGCAAGAGAAAGAGAA 3793
Db 1261 GAAGGAGCTTGGAGAAACAAATTAAGGCGAGTTGTGAGGCGCTCTCGCAAGAGAAAGAGAA 1320
QY 3794 GTGCAAAACACTCCACGGGAGAGCGCTTCCGTGGAAGCGGATCCCAAGTTACCCAGTAG 3853
Db 1321 GTGCAAAACACTCCACGGGAGAGCGCTTCCGTGGAAGCGGATCCCAAGTTACCCAGTAG 1380
QY 3854 CAAGGAGAGCCCACTCACTGCTCGAGTGCAGCAAGCTTTTCAGAACCTTACCACAGCT 3913
Db 1381 CAAGGAGAGCCCACTCACTGCTCGAGTGCAGCAAGCTTTTCAGAACCTTACCACAGCT 1440
QY 3914 GGTCTTGCATCTCAGGCTTCCACAAAGGACCGGAGGCGCGGAGTTCGCCACCAT 3973
Db 1441 GGTCTTGCATCTCAGGCTTCCACAAAGGACCGGAGGCGCGGAGTTCGCCACCAT 1500
QY 3974 GTCTGTGAGGAGGAGCAGCGGGGACGTGTTCTCTGACACTGCGCGCCCTCTGATGA 4033
Db 1501 GTCTGTGAGGAGGAGCAGCGGGGACGTGTTCTCTGACACTGCGCGCCCTCTGATGA 1560
QY 4034 AAATGAGCCGTGATTCGAGGGAAGGTGTTCTGAAGCGGATCTGAGGATGGGCTTCC 4093
Db 1561 AAATGAGCCGTGATTCGAGGGAAGGTGTTCTGAAGCGGATCTGAGGATGGGCTTCC 1620
QY 4094 CGAAGGAATCCATCTGGGTAA 4114
Db 1621 CGAAGGAATCCATCTGGATAA 1641
```

```
RESULT 11
BD085734
LOCUS BD085734 3186 bp DNA linear PAT 27-AUG-2002
DEFINITION Genes from the 20q13 amplicon and their uses.
ACCESSION BD085734
VERSION BD085734.1 GI:22631344
KEYWORDS JP 2001524802-A/10.
```

```
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE other sequences; artificial sequences.
AUTHORS 1 (bases 1 to 3186)
Gray,J.W., Collins,C.C., Hwang,S.I., Godfrey,T., Kowbel,D. and
Rommens,J.
TITLE Genes from the 20q13 amplicon and their uses
JOURNAL Patent: JP 2001524802-A 10 04-DEC-2001;
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
COMMENT OS Artificial Sequence
PN JP 2001524802-A/10
PD 04-DEC-2001
PF 15-JUL-1997 JP 1998506264
PR 15-JUL-1996 US 08/680395,16-OCT-1996 US 08/731499 PR
17-JAN-1997 US 08/785532
PI JOE W GRAY, COLIN CONRAD COLLINS, SOO IN HWANG, TONY GODFREY, PI
DAVID KOWBEL,
PI JOHANNA ROMMENS
PC C12N15/11.C12Q1/68.A61K48/00
CC Description of Artificial Sequence:ZABC1 Open Reading Frame FH
Key Location/Qualifiers
FT source 1..3186
FT /organism='Artificial Sequence'.
FEATURES
source
1..3186
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
ORIGIN
Query Match 15.6%; Score 1559.6; DB 6; Length 3186;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1565; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 8286 AGGTGAAAAACCATCAAAATGTGAATTTGTGAATATGTCGAGCCGAGACATCTCT 8345
Db 1482 AGGTGAAAAACCATCAAAATGTGAATTTGTGAATATGTCGAGCCGAGACATCTCT 1541
QY 8346 GAGGTATCACTTGGAGAGACATCAAGGAAAAACCAACCGATGTTGCTGCAAGTCAA 8405
Db 1542 GAGGTATCACTTGGAGAGACATCAAGGAAAAACCAACCGATGTTGCTGCAAGTCAA 1601
QY 8406 GAACGATGTAAAAATCAGGACATCAAGATGACATTAATTAACCGCTGACAGTCCGCAAC 8465
Db 1602 GAACGATGTAAAAATCAGGACATCAAGATGACATTAATTAACCGCTGACAGTCCGCAAC 1661
QY 8466 CAAAAATTTGAAAAAGATTTTGTGAGTGCCTCAAGATGTTACAGGAGTCCAGCTGCAAA 8525
Db 1662 CAAAAATTTGAAAAAGATTTTGTGAGTGCCTCAAGATGTTACAGGAGTCCAGCTGCAAA 1721
QY 8526 GCAGCTTAAGGAGATGCTTCTGTTTTCAGAAATGTTCTGGGAGCGCTGCTCTCACC 8585
Db 1722 GCAGCTTAAGGAGATGCTTCTGTTTTCAGAAATGTTCTGGGAGCGCTGCTCTCACC 1781
QY 8586 AGCACACAAAGATATCTCAGGATTTCCATAAAAAATGACGCTGATGACAGTCTGATAAAGT 8645
Db 1782 AGCACACAAAGATATCTCAGGATTTCCATAAAAAATGACGCTGATGACAGTCTGATAAAGT 1841
QY 8646 GAATAAAAACCTTACCCCTGCTTACCTGGAATCTGTTTAAAAAGAGATGACAGTGTGAAC 8705
Db 1842 GAATAAAAACCTTACCCCTGCTTACCTGGAATCTGTTTAAAAAGAGATGACAGTGTGAAC 1901
QY 8706 TCAGGCAATACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCGATGGCAG 8765
Db 1902 TCAGGCAATACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCGATGGCAG 1961
QY 8766 TACCACCCATACCTTGAAGTTAGCCCCCAAGAGAGCAAAACGAGACCGCAGCTGCTG 8825
Db 1962 TACCACCCATACCTTGAAGTTAGCCCCCAAGAGAGCAAAACGAGACCGCAGCTGCTG 2021
QY 8826 CAGATACAGGCAAGTGTGATTTGTACGAAAAACCTTTAAATTTATCCGTGGGGCTCT 8885
Db 2022 CAGATACAGGCAAGTGTGATTTGTACGAAAAACCTTTAAATTTATCCGTGGGGCTCT 2081
```

QY	8886	TCACAAATGCCCGGCAATTTCTTTGAGTAAAGTTTGATTCACAGTATCACCTGTCCATT	8945
Db	2082	TCACAAATGCCCGGCAATTTCTTTGAGTAAAGTTTGATTCACAGTATCACCTGTCCATT	2141
QY	8946	TTGTACCTTCAAGACATTTTATCCAGAGTTTATATGATGACACAGAGCTGAGCATAA	9005
Db	2142	TTGTACCTTCAAGACATTTTATCCAGAGTTTATATGATGACACAGAGCTGAGCATAA	2201
QY	9006	ATACAATCCTGACGTTTCATAAAACTGTCGAAACAGTCTTGTCTTAGAGTGCAGTAC	9065
Db	2202	ATACAATCCTGACGTTTCATAAAACTGTCGAAACAGTCTTGTCTTAGAGTGCAGTAC	2261
QY	9066	CGGATGCCCGCCAGGCTGCTCGGAAAAGATGCTCCCTCCCTAGTTCTGTAAACC	9125
Db	2262	CGGATGCCCGCCAGGCTGCTCGGAAAAGATGCTCCCTCCCTAGTTCTGTAAACC	2321
QY	9126	CAAGCCCAAGTCTGCTTTCCCGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGAAGCA	9185
Db	2322	CAAGCCCAAGTCTGCTTTCCCGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGAAGCA	2381
QY	9186	GAGCCCTCTCGGCCAGGCAAGCCCTCTGACTTCAGGATAGACTCTAGCACTTTAGC	9245
Db	2382	GAGCCCTCTCGGCCAGGCAAGCCCTCTGACTTCAGGATAGACTCTAGCACTTTAGC	2441
QY	9246	CCCAAGTAACCTGAAGTCCACAGACACAGCAGAAATGCGGGTCCAAAGGGCCGCCAC	9305
Db	2442	CCCAAGTAACCTGAAGTCCACAGACACAGCAGAAATGCGGGTCCAAAGGGCCGCCAC	2501
QY	9306	CAGGCAACAGCAATCTGAGATGTTTCTTAAACACAGTGTTCCTGACCGGATAAGAC	9365
Db	2502	CAGGCAACAGCAATCTGAGATGTTTCTTAAACACAGTGTTCCTGACCGGATAAGAC	2561
QY	9366	AAAAAGACCCGAGACAAAATTTGAAACCTCTTCAGTAGTCTCTTCAGCCCAACCTCGG	9425
Db	2562	AAAAAGACCCGAGACAAAATTTGAAACCTCTTCAGTAGTCTCTTCAGCCCAACCTCGG	2621
QY	9426	GAGCAGTAACATATGTTTCATGATACATACCCCGCAAGAACAGCAGCCGCTGGGAC	9485
Db	2622	GAGCAGTAACATATGTTTCATGATACATACCCCGCAAGAACAGCAGCCGCTGGGAC	2681
QY	9486	TCCGGGAAGAGACTATTTCTGTAATCGGAGTCCAGCAATCTGACAGCAAAATTTGGTGA	9545
Db	2682	TCCGGGAAGAGACTATTTCTGTAATCGGAGTCCAGCAATCTGACAGCAAAATTTGGTGA	2741
QY	9546	GCCCTTCCAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACAGCCCGGGGC	9605
Db	2742	GCCCTTCCAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACAGCCCGGGGC	2801
QY	9606	CAATTACAGAGAGGCTATGACCTTCCCAAGTACCAATATGGTTCAGAGGATCACATCACT	9665
Db	2802	CAATTACAGAGAGGCTATGACCTTCCCAAGTACCAATATGGTTCAGAGGATCACATCACT	2861
QY	9666	GTTACCGGAGGACTGTGTATCCGTCGAGCGGCTGCTCCCAACCAAGGTTCCCTGAG	9725
Db	2862	GTTACCGGAGGACTGTGTATCCGTCGAGCGGCTGCTCCCAACCAAGGTTCCCTGAG	2921
QY	9726	CTCCAGCAGAGGCTGATTTCTCCAAATGCTGACTGTTTCAGAAAGCCCTATGGTCCCG	9785
Db	2922	CTCCAGCAGAGGCTGATTTCTCCAAATGCTGACTGTTTCAGAAAGCCCTATGGTCCCG	2981
QY	9786	GCCACTTTACATTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	9845
Db	2982	GCCACTTTACATTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3041
QY	9846	TGCATGAGGGCGT	9859
Db	3042	TGTTGATGTCAGT	3055

RESULT 12
AR070326
LOCUS

AR070326 1507 bp DNA linear PAT 18-FEB-2000

DEFINITION Sequence 3 from patent US 5892010.
ACCESSION AR070326
VERSION AR070326.1 GI:7221214
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1507)
AUTHORS Gray, J., Collins, C., Hwang, S.-i., Godfrey, T., Kowbel, D. and
TITLE Genes from the 20Q13 amplicon and their uses
JOURNAL Patent: US 5892010-A 3 06-APR-1999;
FEATURES
source 1..1507
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 13.9%; Score 1387.2; DB 6; Length 1507;
Best Local Similarity 97.3%; Pred. No. 4.7e-311;
Matches 1470; Conservative 3; Mismatches 32; Indels 6; Gaps 6;

QY	2424	CAGGTGCTGGGATGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAAGAGA	2483
Db	2	CAGGTGCTGGGATGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAAGAGC	61
QY	2484	ACTAATGCTTGTGCTGATTCATATTTGAATCGAGGCAATGGGAAACCTGTATGCTTGT	2543
Db	62	ACTAATGCTTGTGCTGATTCATATTTGAATCGAGGCAATGGGAAACCTGTATGCTTGT	121
QY	2544	TTGTGGAAGACCAAGTCACACCACTCAGTCTTCTTAAAGTTTGAAGTTGAGG	2603
Db	122	TTGTGGAAGAACCAAGTCACACCACTCAGTCTTCTTAAAGTTTGAAGTTGAGG	181
QY	2604	ACTATACACTTCTTTTGAACCTTTTATATAAATATTTGCTCTGTTTGGAAACCCAGG	2663
Db	182	ACTATACACTTCTTTTGAACCTTTTATATAAATATTTGCTCTG- TTTTGAACCCAGG	240
QY	2664	GCTGTTAGAGGGGTGAGTGACAAAGTCTTACAAAGTGGCTTATTCACACTCCAGAAATGC	2723
Db	241	ACTGTTAGA- GGGTGAAGTGAAGTCTTAC- AGTGGCTTAACTCAACTCCAGAAATGC	298
QY	2724	CCACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCAACTCAATCC	2783
Db	299	CCACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCAACTCAATCC	358
QY	2784	CTCTTAATGTACATGGATGGCCAGAAAGTATGGCAGCTCTCTTGGCAGTCCGATGGAG	2843
Db	359	CTCTTAATGTACATGGAT- GGCCAAAGTATGGCAGCTCTCTTGGCAGTCCGATGGAG	417
QY	2844	ATGGAGGATGCTTGTCAATGAAAGGACCGCTGTTGTTCCATTCGAGCTACACAAGAA	2903
Db	418	ATGGA- GATGCTTGTCAATGAAAGGACCGCTGTTGTTCCATTCGAGCTACACAAGAA	476
QY	2904	AAAAATGTC- ATCCAAATCGAGGGTATATGCTTGGATTCATGCTTCTGAGCCAGAC	2962
Db	477	AAAAATGTCATTCGAAATCGAGGGTATATGCTTGGATTCATGCTTCTGAGCCAGAC	536
QY	2963	CTTACACATTCAGAAAGCTTTAAATAACATGCTTTAAATGCAACACCGGCTTACCCTCTG	3022
Db	537	CTTACACATTCAGAAAGCTTTAAATAACATGCTTTAAATGCAACACCGGCTTACCCTCTG	596
QY	3023	TGAACAGAGGTTCTTGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGT	3082
Db	597	TGAACAGAGGTTCTTGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGT	656
QY	3083	CGGACAGAACTCCCAAGGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGTATG	3142
Db	657	CGGACAGAACTCCCAAGGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGTATG	716
QY	3143	TGGGAGACATTTAGAGTCCGCTTTTGTGATGTTGAGATCCACATGAGAAACACAAAGATT	3202
Db	717	TGGGAGACATTTAGAGTCCGCTTTTGTGATGTTGAGATCCACATGAGAAACACAAAGATT	776

```
QY 3203 TTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAA 3262
DB 777 TTTCACTTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAA 836
QY 3263 TCACATCGGCACACATATATGCGAAATCGGGGGCCAGAAGCAAACTGCAGCAAGCCTTGGGA 3322
DB 837 TCACATCGGCACACATATGCGAAATCGGGGGCCAGAAGCAAACTGCAGCAAGCCTTGGGA 896
QY 3323 GAGTAGTCCAGCAACGATCAACGAGTCCGTCAGGTGCGCGCGCGAGAGCATCTCCTC 3382
DB 897 GAGTAGTCCAGCAACGATCAACGAGTCCGTCAGGTGCGCGCGCGAGAGCATCTCCTC 956
QY 3383 TCCTTACAAAATCTGCATGGTTTGTGGCTTCTCTATTTCCAAATAAGAAAGTCTTAATTGA 3442
DB 957 TCCTTGCAAAATCTGCATGGTTTGTGGCTTCTCTATTTCCAAATAAGAAAGTCTTAATTGA 1016
QY 3443 GCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACAGAGCGCGCGAGACAGACTC 3502
DB 1017 GCACCGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACAGAGCGCGCGAGACAGACTC 1076
QY 3503 TCCACAAGGAGGAATGCGCTCCTCGAGGAGGACTTCTCGCAGTTGTTCAACTTGGAGACC 3562
DB 1077 TCCACAAGGAGGAATGCGCTCCTCGAGGAGGACTTCTCGCAGTTGTTCAACTTGGAGACC 1136
QY 3563 AAAATCTCACCTCGAAACGGGGGAAGAGCCTGTGCAGATGCATCCCTCAGCTCGATCCGTT 3622
DB 1137 AAAATCTCACCTCGAAACGGGGGAAGAGCCTGTGCAGATGCATCCCTCAGCTCGATCCGTT 1196
QY 3623 CACACCTTCCAGCTTGGCAGCTGCTACCAAGGAAGAGTTGCCATTTGCCAAGAGT 3682
DB 1197 CACACCTTCCAGCTTGGCAGCTGCTACCAAGGAAGAGTTGCCATTTGCCAAGAGT 1256
QY 3683 GAAGGAATCGGGCAAGAGGAGCAGCAGACAGCAGATTCCGATTCGAGTCCGAGAGGAGCT 3742
DB 1257 GAAGGAATTCGGGCAAGAGGAGCAGCAGACAGCAGATTCCGATTCGAGTCCGAGAGGAGCT 1316
QY 3743 TGGAGAAAACAAATAAGGCGCAGTTGTGTCAGGCGCTCTCGCAAGAGAAAAGAGTGCAAAACA 3802
DB 1317 TGGAGAAAACAAATAAGAACCAATTGTGTCAGGCGCTCTCGCAAGAGAAAAGAGTGCAAAACA 1376
QY 3803 CTCGCCAGGGAAGCGCCTCCGTGAGCGCGGATCCCAAGTTACCAAGTAGCAAGAGAGAA 3862
DB 1377 CTCGCCAGGGAAGCGCCTCCGTGAGCGCGGATCCCAAGTTACCAAGTAGCAAGAGAGAA 1436
QY 3863 GCCCACTCACTGCTCCGAGTGCAGCAAGCTTTCAGAACCTACCAAGCTGCTGCTTGA 3922
DB 1437 GCCCACTCACTGCTCCGAGTGCAGCAAGCTTTCAGAACCTACCAAGCTGCTGCTTGA 1496
QY 3923 CTCAGGGGTCC 3933
DB 1497 CTCAGGGGTCC 1507

RESULT 13
BD085727
LOCUS Genes from the 20q13 amplicon and their uses.
DEFINITION 1507 bp DNA linear PAT 27-AUG-2002
ACCESSION BD085727
VERSION BD085727.1 GI:22631337
KEYWORDS JP 2001524802-A/3.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 1507)
AUTHORS Gray,J.W., Collins,C.C., Hwang,S.I., Godfrey,T., Kowbel,D. and Rommens,J.
TITLE Genes from the 20q13 amplicon and their uses
JOURNAL Patent: JP 2001524802-A 3 04-DEC-2001;
COMMENT THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
OS Artificial Sequence
PN JP 2001524802-A/3
PD 04-DEC-2001
```

```
PF 15-JUL-1997 JP 1998506264
PR 15-JUL-1996 US 08/680395,16-OCT-1996 US 08/731499 PR
17-JAN-1997 US 08/785532
PI JOE W GRAY, COLIN CONRAD COLLINS, SOO IN HWANG, TONY GODFREY, PI
DAVID KOWBEL,
PI JOHANNA ROMMENS
PC C12N15/11.C12Q1/68.A61K48/00
CC Description of Artificial Sequence:cc49
CC N is A, G, T, or U
FH Key Location/Qualifiers
FT modified base (447).
FEATURES
    source
        1..1507
            /organism="synthetic construct"
            /mol_type="genomic DNA"
            /db_xref="taxon:32630"
ORIGIN
Query Match 13.9%; Score 1387.2; DB 6; Length 1507;
Best Local Similarity 97.3%; Pred. No. 4.7e-311;
Matches 1470; Conservative 3; Mismatches 32; Indels 6; Gaps 6;
QY 2424 CAGGTTGCTGGGATTCGACTTCTTGCTCAATTGAAACACTCAATTCGAGAGACAAAGAGA 2483
DB 2 CAGGTTGCTGGGATTCGACTTCTTGCTCAATTGAAACACTCAATTCGAGAGACAAAGAGC 61
QY 2484 ACTAATGCTTGTGCTGATTTCATATTTGAATCGAGGCATTTGGAAACCTGTATGCCCTGT 2543
DB 62 ACTAATGCTTGTGCTGATTTCATATTTGAATCGAGGCATTTGGAAACCTGTATGCCCTGT 121
QY 2544 TTGTGAAAAGAACCACTGACCACTCACTGAGCTTCTCTAAAAGTTTCCGAAAGTTAGAGG 2603
DB 122 TTGTGAAAAGAACCACTGACCACTCACTGAGCTTCTCTAAAAGTTTCCGAAAGTTAGAGG 181
QY 2604 ACTATACATTTCTTTTGAACCTTTTATAATAATATTTGCTCTGGTTTTTGGAAACCCAGG 2663
DB 182 ACTATACATTTCTTTTGAACCTTTTATAATAATATTTGCTCTGG-TTTTGGAAACCCAGG 240
QY 2664 GCTGTTAGAGGGTGAGTCAGCAAGCTTCAAGTGGCCTTATTTCCAACTCCAGAAATTGC 2723
DB 241 ACTGTTTAGA-GGTGAGTGACAGGCTCTTAC-AGTGGCCTTTAATCCAACTCCAGAAATTGC 298
QY 2724 CCAACGGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATTCGCAATCC 2783
DB 299 CCAACGGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATTCGCAATCC 358
QY 2784 CTCTTAATGTACATGATGGCCAGAAAGTGAATGGCAGCTCTCTTGGCAGTCCGATGGAG 2843
DB 359 CTCTTAATGTACATGATGGAT-GGCCAAGAGTGAATGGCAGCTCTCTTGGCAGTCCGATGGAG 417
QY 2844 ATGGAGGATGCCTTGTCAATGAAGGACCGCTGTTGTTCCATTCCGAGCTACACAAGAA 2903
DB 418 ATGGA-GATGCTTGTCAATGAAGGACCGCTGTTGTTGTTGTTCAATTCGAGCTACACAAGAA 476
QY 2904 AAAAAATGTC-ATCCAAATCGAGGGTATATGCGCTTGGATTGCAATGTTCTGACGCCAGAC 2962
DB 477 AAAAAATGTC-ATCCAAATCGAGGGTATATGCGCTTGGATTGCAATGTTCTGACGCCAGAC 536
QY 2963 CTTCACACATTCAGAAAGCCTTAAATAACAATGCTTTAAATGCAACACCGGCTTACCTCTG 3022
DB 537 CTTCACACATTCAGAAAGCCTTAAATAACAATGCTTTAAATGCAACACCGGCTTACCTCTG 596
QY 3023 TGAACCAAGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATATAAGTCAAGT 3082
DB 597 TGAACCAAGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATATAAGTCAAGT 656
QY 3083 GCGAACAGAACCTTCCCAAGGAAAAGAAATTCGAAGGAAAATGAATTTAGCTGTGAGGTATG 3142
DB 657 GCGAACAGAACCTTCCCAAGGAAAAGAAATTCGAAGGAAAATGAATTTAGCTGTGAGGTATG 716
QY 3143 TGGGCAGACATTTAGAGTCCGCTTTTGAATGTTGAGATCCCATGAGAAACACACAAAGATTC 3202
DB 717 TGGGCAGACATTTAGAGTCCGCTTTTGAATGTTGAGATCCCATGAGAAACACACAAAGATTC 776
```

```
QY 3203 TTTCACTTACGGGTGTACATGTCGAGAGAGATTCAAGGAGCCTTGGTTCTTAAAAA 3262
Db 777 TTTCACTTACGGGTGTACATGTCGAGAGAGATTCAAGGAGCCTTGGTTCTTAAAAA 836
QY 3263 TCACATGCGGACACATAATGGCAATCGGGGCCAGAGCAAACTGCAGCAAGGCTTGA 3322
Db 837 TCACATGCGGACACATAATGGCAATCGGGGCCAGAGCAAACTGCAGCAAGGCTTGA 896
QY 3323 GAGTAGTCCAGCAACGATCAACGAGGTGTCAGGTGTCAGCGGCGGAGAGCATCTCCTC 3382
Db 897 GAGTAGTCCAGCAACGATCAACGAGGTGTCAGGTGTCAGCGGCGGAGAGCATCTCCTC 956
QY 3383 TCCTTACAAATCTGCATGGTTTGGCTTCTCTATTCCAAATAAGAAAGTCTAAATTGA 3442
Db 957 TCCTTGCAAAATCTGCATGGTTTGGCTTCTCTATTCCAAATAAGAAAGTCTAAATTGA 1016
QY 3443 GCACCGCAGGTGCACCAAAAAAATGCTTTTCGGTACGACGCGGCGAGAGACTC 3502
Db 1017 GCACCGCAGGTGCACCAAAAAAATGCTTTTCGGTACGACGCGGCGAGAGACTC 1076
QY 3503 TCCACAAGGAGGAATGCCGTCTCGAGGAGGACTTCTCGCAGTTGTTTCAACTTGAGACC 3562
Db 1077 TCCACAAGGAGGAATGCCGTCTCGAGGAGGACTTCTCGCAGTTGTTTCAACTTGAGACC 1136
QY 3563 AAAATCTCACCTGAAACGGGGAAGAGCTGTGATGATCCTCCTCAGCTCGATCCGTT 3622
Db 1137 AAAATCTCACCTGAAACGGGGAAGAGCTGTGATGATCCTCCTCAGCTCGATCCGTT 1196
QY 3623 CACCACTTCCAGGCTTGGCAGCTGGCTACCAAGGAAAGTTGCCATTGCAAGAAAGT 3682
Db 1197 CACCACTTCCAGGCTTGGCAGCTGGCTACCAAGGAAAGTTGCCATTGCAAGAAAGT 1256
QY 3683 GAAGGAATCGGGGCAAGAGGAGCAACGACACGATTCGAGTTCGAGAGGAGCT 3742
Db 1257 GAAGGAATCGGGGCAAGAGGAGCAACGACACGATTCGAGTTCGAGAGGAGCT 1316
QY 3743 TGGAGAAACAAATAAGGCGAGTTGTCAGGCTCTCGCAGAGAAAGAGAGTGCACACA 3802
Db 1317 TGGAGAAACAAATAAGGCGAGTTGTCAGGCTCTCGCAGAGAAAGAGAGTGCACACA 1376
QY 3803 CTCACAGCGGAGGCGCTCCGTCGAGCGCGATCCAGTTACCGAGTACGACGAGAA 3862
Db 1377 CTCACAGCGGAGGCGCTCCGTCGAGCGCGATCCAGTTACCGAGTACGACGAGAA 1436
QY 3863 GCCCACTCACTGCTCGAGTGGCGCAAGCTTTTCAGAACCTTACCAACAGCTGGTTGCA 3922
Db 1437 GCCCACTCACTGCTCGAGTGGCGCAAGCTTTTCAGAACCTTACCAACAGCTGGTTGCA 1496
QY 3923 CTCACAGGCTCC 3933
Db 1497 CTCACAGGCTCC 1507
```

```
RESULT 14
LOCUS HS449017/c
DEFINITION Human DNA sequence from clone RP3-449017 on chromosome
22q13.1-13.2, complete sequence.
ACCESSION Z98749
VERSION Z98749.11 GI:4775611
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1. (bases 1 to 59586)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Lloyd, D.
Direct Submission
TITLE Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT On May 11, 1999 this sequence version replaced gi:4688875.
```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/c_elegans/wormpep -----

Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: humquery@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

This sequence was generated from part of bacterial clone contigs of human chromosome 22, constructed by the Sanger Centre Chromosome 22 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr22>

RP3-449017 is from the library RP3-3 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2.

FEATURES

Location/Qualifiers
1..59586
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="22"
/map="q13.1-13.2"
/clone="RP3-449017"
/clone_lib="RP3-3"

ORIGIN

Query Match 12.2%; Score 1223.6; DB 9; Length 59586;
Best Local Similarity 97.8%; Pred. No. 8.4e-273;
Matches 1264; Conservative 0; Mismatches 19; Indels 9; Gaps 2;
QY 1027 TCCATTGTGATAGCAACAAAGAAAGACTTCAGGACATTCTTCTTACCCCTAATCCTT 1086
Db 59586 TCCATTGTGATAGCAACAAAGAAAGACTTCAGGACATTCTTCTTACCCCTAATCCTT 59527
QY 1087 GATCTGCACTCTTATTAGAAAAGCTTAATGTTAAAGATCTAGTTATTCAAAACTAAAG 1146
Db 59526 GATCTGCACTCTTATTAGAAAAGCTTAATGTTAAAGATCTAGTTATTCAAAACTAAAG 59467
QY 1147 ATACAAGAGATATGAGAAATTTCTATTTCGGAGTTAAAGGAGAGATGTTTCTTGGCT 1206
Db 59466 ATACAAGAGATATGAGAAATTTCTATTTCGGAGTTAAAGGAGAGATGTTTCTTGGCT 59407
QY 1207 TCTCTGAGCTGCGAGGCTTCTCTCTCTTTAAAGAAAGTAGAGAGAGGAGGAGATAA 1266
Db 59406 TCTCTGAGCTGCGAGGCTTCTCTCTCTTTAAAGAAAGTAGAGAGAGGAGGAGATAA 59347
QY 1267 GTATGCTTTTGTGTTTAAAGTTACTTCTGCGGAGTAGTTGTCAGTCTTTTGGTTTC 1326
Db 59346 GTATGCTTTTGTGTTTAAAGTTACTTCTGCGGAGTAGTTGTCATCCCTTTTGGTTTC 59287
QY 1327 TTGGGTGAATTAAGTCTTAAGTTTAAAGTAGTTGGGACTTATTTAAAGAAAGTGCCTTA 1386
Db 59286 TTGGGTGAATTAAGTCTTAAGTTTAAAGTAGTTGGGACTTATTTAGAAACAAAGTGCCTTA 59227
QY 1387 TCCAATGTTTGCCATAAAGGAGAGGGTATTGGCTTTAGAAAGTTAATTTCTTCCAGGAG 1446

Db 59226 TCCAAATGTTGGCCATAAAGCAGAGGGTATTTGGCTTTAGAGTAAATTTCTTCTCCAGGAG 59167
Qy 1447 TGAATAATTAGCTCTTAAACCAAGACGACGAGCTAAATAAGTAATTTTCCACCTGGCC 1506
Db 59166 TGAATAATTAGCTCTTAAACCAAGACGACGAGCTAAATAAGTAATTTTCCACCTGGCC 59107
Qy 1507 AGTGCATGATGTGAAGGTAGATTAAATAATGAGAGGCCATTTTCTGATGAAGACT 1566
Db 59106 AGTGCATGATGTGAAGGTAGATTAAATAATGAGAGGCCATTTTCTGATGAAGACT 59047
Qy 1567 AAGCCATGTTGAACAGCCCTGTGAGGATTTTATTTAAATCTATACATTCACAAGGA 1526
Db 59046 AAGCCATGTTGAACAGCCCTGTGAGGATTTTATTTAAATCTATACATTCACAAGGA 59987
Qy 1627 GCTTTGTATGATCTTTCCCTATTTTGTGTTGAGCTAGGAAGCCACCCAGCTGCTTGT 1686
Db 59986 GCTTTGTATGATCTTTCCCTATTTTGTGTTGAGCTAGGAAGCCACCCAGCTGCTTGT 59927
Qy 1687 TGAAGCAGAAAGTCTGTGAAGCAAGCTGGGATTTTGAACAGTGGATTGAGGTTTCGAAT 1746
Db 59926 TGAAGCAGAAAGTCTGTGAAGCAAGCTGGGATTTTGAACAGTGGATTGAGGTTTCGAAT 59867
Qy 1747 ATCCAGTGAACCAAAATATATACAGGTTCCCTGGCCAGACGATGACCATTTCTGAGGT 1806
Db 59866 ATCCAGTGAACCAAAATATATACAGGTTCCCTGGCCAGACGATGACCATTTCTGAGGT 59807
Qy 1807 GTTAAGTATTTCTGTAATGGGATTTTAGGAAGTTTCTGTATTTCTGTGCTCATTTTG 1866
Db 59806 GTTAAGTATTTCTGTAATGGGATTTTAGGAAGTTTCTGTATTTCTGTGCTCATTTTG 59747
Qy 1867 TTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTGTTGGGCACTTTAGATTTCTTGATG 1926
Db 59746 TTGACCTCTGTATGTGCAAAATCTCTAAGGGGGTGTGTTGGGCACTTTAGGATTTCTGATG 59691
Qy 1927 CAGATTTGTTGATATGAACAAATTTTAAATGTTTGTATATACACTGATGATTTTAAAA 1984
Db 59690 CAGATTTGTTGATATGAACAAATTTTAAATGTTTGTATATACACTGATGATTTTAAAA 59631
Qy 1985 ---TAGTTTACTAAAGTGTTTTAAATTTTTCATCTTAAATTTTTCACAGTCTTATATGCTTT 2041
Db 59630 TTTTAGTTTACTAAAGTGTTTTAAATGTTTTCATCTTAAATTTTTCACAGTCTTATATGCTTT 58571
Qy 2042 TAGATTTAGGGAGGCTGTTGATGGCATCCACATGTCATTTTAGTGGCATTTTAAAAATGTA 2101
Db 59570 TAGATTTAGGGAGGCTGTTGATGGCATCCACATGTCATTTTAGTGGCATTTTAAAAATGTA 58511
Qy 2102 TTCAGCTGAATTTTAAACATTTCTGACCTAAACCTTGACATTTTAGATTTTAAAGTCGGTAAA 2161
Db 59510 TTCAGCTGAATTTTAAACATTTCTGACCTAAACCTTGACATTTTAAAGTCGGTAAA 58451
Qy 2162 GCACCTGATTTAACTGGATTTTAACTGGATGAAATTTCTGATTTTAAATGTTACTGACTG 2221
Db 59450 GCACCTGATTTAACTGGATTTTAACTGGATGAAATTTCTGATTTTAAATGTTACTGACTG 58391
Qy 2222 GATAAATGCCAATGATTTTAAATTAACCAAGCAGCTTTTAAAGGATGCTTATATATAGTT 2281
Db 59390 GATAAATGCCAATGATTTTAAATTAACCAAGCAGCTTTTAAAGGATGCTTATATATAGTT 58331
Qy 2282 AAGAGTGAAGCAATTTGATGATGCTTCTC 2313
Db 59330 AAGAGTGAAGCAATTTGATGATGCTTCTC 58299

RESULT 15
HS434P1/c 124990 bp DNA linear PRI 04-MAR-2003
LOCUS Human DNA sequence from clone RP3-434P1 on chromosome 22, complete
DEFINITION sequence.
ACCESSION 297056
VERSION 297056.1 GI:2832593
KEYWORDS HTG.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 124990)
Conor, R.
Direct Submission
Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Feb 5, 1998 this sequence version replaced gi:2462403.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep -----
Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
This sequence was generated from part of bacterial clone contigs of
human chromosome 22, constructed by the Sanger Centre Chromosome 22
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr22
RP3-434P1 is from the library RP3-3 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2.
Location/Qualifiers
1..124990
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="RZPD:RP3-434P1" 1434
/db_xref="taxon:9606"
/chromosome="22"
/clone="RP3-434P1"
/clone_lib="RP3-3"
ORIGIN
Query Match 10.1%; Score 1006.6; DB 9; Length 124990;
Best Local Similarity 95.5%; Pred. No. 2,3e-222;
Matches 1083; Conservative 0; Mismatches 37; Indels 14; Gaps 4;
Qy 1 CCATCATATTTCTTATTTTGGCGGAGAGGG-----AGCTTGTCTGTGTCGCCA 54
Db 1128 CCATCATATTTCTTATTTTGGCGGAGAGGGAGTCAAAAGTCTCGTCTGTGTCGCCA 1069
Qy 55 GGCTGGA--CCAGTGGTGGATCTTGGCTCACTGCAACCTCCACCTCCCTGGGTTCAAGTG 112
Db 1068 GGCTGGAATGCAAGTGGCGGATCTTGGCTCACTGCAACCTCCACCTCCCTGGGTTCAAGTG 1009
Qy 113 ATTCCCAATAGCTGGGATTCAGAGTGTGTATTACCATGCCAGCTAAATTTTGTATTTT 172
Db 1008 ATTCCCAAGTACCTGGGATTCAGAGTGTGTATTACCATGCCAGCTAAATTTTGTATTTT 949
Qy 173 TAGCAGATAAGGGGTTTACCATGTTGGCCAGGCTGCTCAACCTCTGCTGCTCATGTG 232
Db 948 TAGCAGATAAGGGGTTTACCATGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 889

QY 233 ATCCACCCACTTCGGCTTCCCAAAAGCATTTGGAGTATAGGTGTGAGCCACTATACCCGTC 292
Db |||||||
QY 888 ATCCACCCACTTCGGCTTCCCAAAAGCATTTGGAGTACAGGTGTGAGCCACTATGCCCCGC 829
Db |||||||
QY 293 CTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTCTAAAGGATGT 352
Db |||||||
QY 828 CTCACATCATATTTCTAATCCCGAGACTGTAGAGCTGGTGTCTCTTTTGGAAAGGATGT 769
Db |||||||
QY 353 CAGTAGAGAAGTGGAGTTCCTCCCAAAATTAACGTTTTCACGTAATAGTCAAGTTTCTAAAT 412
Db |||||||
QY 768 CAGTAGAGAAGTGGAGTTCCTCCCAAAATTAACGTTTTCACATATAGTCAAGTTTCTAAAT 709
Db |||||||
QY 413 ACAGTAATAATGTTGAGAGCTGACATAGGACTAACTTGGTGTCTTTTCTTTTCTTTTCTTTT 472
Db |||||||
QY 708 ACAGTAATAATGTTGAGAGCTGACGTAAGGACTAACTTAG-----TTTTTGTCTTTTCTTT 654
Db |||||||
QY 473 TTTCAAAATTCACCTGAACCTTTGATTTTCTAAATAAGGACATTAATAAAAAAACCAAAA 532
Db |||||||
QY 653 TTTCAAAATTCACCTGAACCTTTGATTTTCTAAATAAGGACATTTAAATAAACCAAAA 595
Db |||||||
QY 533 AACTCCACTATTGCCCTATTGCCACTATTGATTTTAAATAAAGCGTATTTTAGCAT 592
Db |||||||
QY 594 AACTCCACTATTGCCCTATTGCCACTATTGATTTTAAATAAAGCGTATTTTAGCAT 535
Db |||||||
QY 593 CTAAAGTAGGAAGGACCTCAAAATAAATGAGTCTTTTCTTTGGCCAGGGAACACAGCGT 652
Db |||||||
QY 534 CTGCAAGTAGGAAGGCGCTCAAAATAAATGAGTCTTTTCTTTGGCCAGGGAACACAGCGT 475
Db |||||||
QY 653 TGTGAGAAATTCATTAACCTTTTCTAGGGTATGCTGTATTCAGTTTAAACCTTGCC 712
Db |||||||
QY 474 TGTGAGCAATTTGTAACCTTTTCTAGGGTATGCTGTATTCAGTTTAAACCTTGCC 415
Db |||||||
QY 713 TGGGAGCGCTAGCAATTCAGTAAATACCTTGTGAATAAGCAAAATGAACTTAAAGCTTCTATG 772
Db |||||||
QY 414 TGGGAGCGCTAGCAATTCAGTAAATACCTTGTGAATAAGCAAAATGAACTTAAAGCTTCTATG 355
Db |||||||
QY 773 TATAGAAACCTAAGTCACTTTCACATTCGATTTAGCAGAGTAATGAAATTTCTTTTCAAT 832
Db |||||||
QY 354 TATAGAAACCTAAGTCACTTTCACATTCGATTTAGCAGAGTAATGAAATTTCTTTTCAGT 295
Db |||||||
QY 833 GTGTAGCTCTATCCCGAAGCCACAGAAATTTGGAACTGTAAAGGCCATCCTATAGTTTA 892
Db |||||||
QY 294 GTGTAGATCTATCCCGAAGCCACAGAAATTTGGAACTGTAAAGGCCATCCTAGAGTTTA 235
Db |||||||
QY 893 ACCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGT 952
Db |||||||
QY 234 ACCAACTGCGTTAAATAGATAATAGAAAGATGTGGTATGTGGCAGTGACAACTTGAAGGT 175
Db |||||||
QY 953 TGTGACTAGAACTCGGGTCTCTGGAGTGTCTATTATATACACCAAGCTGGTCAACGAC 1012
Db |||||||
QY 174 TGTGACTAGAACTCGGGTCTCTGGAGTGTCTATTATATACACCAAGCTGGTCAACGAC 115
Db |||||||
QY 1013 CCATGTGTTGATCTCTCCTGATGTTGATAGCAACAAAGAAAGACTTCAGGACATTTCTTCCT 1072
Db |||||||
QY 114 CCATGTGTTGATCTCTCCTGATGTTGATAGCAACAAAGAAAGACTTCAGGACATTTCTTCCT 55
Db |||||||
QY 1073 TTACCCCTAATCTGATCTGCAAGTCTTATTAGAAAGCTTAAATGTTAAAGATC 1126
Db |||||||
QY 54 TTACCCCTAATCTGATCTTCAGTCTTATTAGAAAGCTTAAATGTTAAAGATC 1

Search completed: July 31, 2005, 20:22:46
Job time : 27419 secs